

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 18:00:47 / Search time 24 Seconds
(without alignments)
30.115 Million cell updates/sec

Title: US-10-045-673A-1
Perfect score: 68
Sequence: 1 TPELAKLVNKRSE 14

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	95.6	89	4	US-08-618-485B-2
2	65	95.6	94	4	US-08-618-485B-3
3	65	95.6	458	4	US-08-618-485B-1
4	65	95.6	458	6	5177002-1
5	59	86.8	458	6	5177002-2
6	59	86.8	474	1	US-08-222-619-5
7	59	86.8	474	5	PCT-US95-04075-5
8	41	60.3	898	4	US-09-585-858-37
9	39	57.4	82	4	US-08-311-731A-351
10	39	57.4	99	4	US-08-311-731A-76
11	39	57.4	115	4	US-08-311-731A-295
12	39	57.4	496	2	US-08-949-637-2
13	39	57.4	496	3	US-09-291-488-2
14	39	57.4	507	4	US-09-153-277-4
15	39	57.4	695	4	US-09-153-277-2
16	38	55.9	556	2	US-08-505-377-1
17	38	55.9	556	3	US-08-798-269-1
18	38	55.9	556	4	US-09-035-210-1
19	38	55.9	556	4	US-09-298-924-8
20	38	55.9	749	4	US-09-489-039A-7979
21	37	54.4	65	4	US-09-134-001C-4977
22	37	54.4	86	2	US-08-809-740A-3
23	36	52.9	250	4	US-09-198-452A-859
24	36	52.9	256	4	US-09-134-001C-3271
25	36	52.9	402	4	US-09-252-991A-20804
26	36	52.9	970	6	5229293-2
27	35.5	52.2	1504	4	US-09-328-352-7046

28	35	51.5	84	4	US-09-543-681A-7825	Sequence 7825, Ap
29	35	51.5	174	4	US-09-107-532A-4968	Sequence 4968, Ap
30	35	51.5	214	3	US-09-587-066-6	Sequence 6, Appl1
31	35	51.5	256	4	US-08-956-171E-5228	Sequence 5228, Ap
32	35	51.5	275	4	US-09-107-532A-4013	Sequence 4013, Ap
33	35	51.5	275	4	US-09-134-000C-6528	Sequence 6528, Ap
34	35	51.5	322	1	US-08-327-494A-2	Sequence 2, Appl1
35	35	51.5	322	5	PCT-US95-13659-2	Sequence 2, Appl1
36	35	51.5	478	4	US-09-134-000C-4712	Sequence 4712, Ap
37	35	51.5	821	1	US-08-339-578-2	Sequence 2, Appl1
38	35	51.5	1160	4	US-09-328-352-6457	Sequence 6457, Ap
39	35	51.5	1525	4	US-09-418-710-69	Sequence 69, Appl
40	35	51.5	1527	4	US-09-418-710-27	Sequence 27, Appl
41	35	51.5	1531	4	US-09-418-710-29	Sequence 29, Appl
42	34	50.0	60	4	US-09-621-976-6141	Sequence 6141, Ap
43	34	50.0	86	4	US-09-621-976-6582	Sequence 6582, Ap
44	34	50.0	145	4	US-09-134-000C-4292	Sequence 4292, Ap
45	34	50.0	243	4	US-09-107-532A-4665	Sequence 4665, Ap

ALIGNMENTS

RESULT 1
US-08-618-485B-2
Sequence 2, Application US/08618485B
Patent No. 6410269
GENERAL INFORMATION:
APPLICANT: NOBITO YAMAMOTO
TITLE OF INVENTION: PREPARATION OF POTENT
MACROPHAGE ACTIVATING FACTORS
TITLE OF INVENTION: DERIVED FROM CLONED VITAMIN D
BINDING PROTEIN AND ITS DOMAIN
TITLE OF INVENTION: AND THEIR THERAPEUTIC USAGE
TITLE OF INVENTION: FOR CANCER, HIV-INFECTION AND
OSTEOPETROSIS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSER: CAESAR, RIVISE, BERNSTEIN,
ADDRESSER: COHEN & FOXOTILOW, LTD.
STREET: 1635 Market Street, 12th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2212
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44 MB
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT VERSION 4.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,485B
FILING DATE: March 19, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,121
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Robert S. Silver
REGISTRATION NUMBER: 35,681
REFERENCE/DOCKET NUMBER: Y1004/20002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2010
TELEFAX: (215) 751-1142
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 89 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: no
ORIGINAL SOURCE:
ORGANISM: Human

```
INDIVIDUAL ISOLATE: Vitamin D-binding protein (Gc protein)
PUBLICATION INFORMATION:
AUTHORS: Cooke, Nancy E., David, E Vivek
TITLE: Serum Vitamin D-binding Protein is a Third Member
of the Albumin and Alpha Fetoprotein Gene Family
JOURNAL: J. Clinical Investigation
VOLUME: 76
ISSUE: 12
PAGES: 2420-2424
DATE: December, 1985
RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 4 and 5 TO 89
US-08-618-485B-2

Query Match          95.6%; Score 65; DB 4; Length 89;
Best Local Similarity 92.9%; Pred. No. 9.4e-05;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPTETLAKLVNKRSE 14
DB 49 TPTETLAKLVNKRSD 62

RESULT 2
US-08-618-485B-3
Sequence 3, Application US/08618485B
Patent No. 6410269
GENERAL INFORMATION:
APPLICANT: NOBUTO YAMAMOTO
TITLE OF INVENTION: PREPARATION OF POTENT
MACROPHAGE ACTIVATING FACTORS
TITLE OF INVENTION: DERIVED FROM CLONED VITAMIN D
BINDING PROTEIN AND ITS DOMAIN
TITLE OF INVENTION: AND THEIR THERAPEUTIC USAGE
TITLE OF INVENTION: FOR CANCER, HIV-INFECTION AND
OSTEOBETROSIS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAESAR, RIVISE, BERNSTEIN,
ADDRESSER: COHEN & POKOTILOW, LTD.
STREET: 1635 Market Street, 12th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2212
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44 MB
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT VERSION 4.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,485B
FILING DATE: March 19, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,121
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Robert S. Silver
REGISTRATION NUMBER: 35,681
REFERENCE/DOCKET NUMBER: Y1004/20002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2010
TELEFAX: (215) 751-1142
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: no
ORIGINAL SOURCE:
ORGANISM: Human
INDIVIDUAL ISOLATE: Vitamin D-binding protein (Gc protein)
```

```
PUBLICATION INFORMATION:
AUTHORS: Cooke, Nancy E., David, E Vivek
TITLE: Serum Vitamin D-binding Protein is a Third Member
of the Albumin and Alpha Fetoprotein Gene Family
JOURNAL: J. Clinical Investigation
VOLUME: 76
ISSUE: 12
PAGES: 2420-2424
DATE: December, 1985
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 10 TO 94
US-08-618-485B-3

Query Match          95.6%; Score 65; DB 4; Length 94;
Best Local Similarity 92.9%; Pred. No. 9.9e-05;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPTETLAKLVNKRSE 14
DB 54 TPTETLAKLVNKRSD 67

RESULT 3
US-08-618-485B-1
Sequence 1, Application US/08618485B
Patent No. 6410269
GENERAL INFORMATION:
APPLICANT: NOBUTO YAMAMOTO
TITLE OF INVENTION: PREPARATION OF POTENT
MACROPHAGE ACTIVATING FACTORS
TITLE OF INVENTION: DERIVED FROM CLONED VITAMIN D
BINDING PROTEIN AND ITS DOMAIN
TITLE OF INVENTION: AND THEIR THERAPEUTIC USAGE
TITLE OF INVENTION: FOR CANCER, HIV-INFECTION AND
OSTEOBETROSIS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAESAR, RIVISE, BERNSTEIN,
ADDRESSER: COHEN & POKOTILOW, LTD.
STREET: 1635 Market Street, 12th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2212
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44 MB
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT VERSION 4.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/618,485B
FILING DATE: March 19, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/478,121
FILING DATE: 07-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Robert S. Silver
REGISTRATION NUMBER: 35,681
REFERENCE/DOCKET NUMBER: Y1004/20002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2010
TELEFAX: (215) 751-1142
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 458 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: no
ORIGINAL SOURCE:
ORGANISM: Human
INDIVIDUAL ISOLATE: Vitamin D-binding protein
INDIVIDUAL ISOLATE: (Gc protein)
```

PUBLICATION INFORMATION:
AUTHORS: Cooke, Nancy E., David, E Vivek
TITLE: Serum Vitamin D-binding Protein is a
TITLE: Third Member of the Albumin and Alpha
TITLE: Fetoprotein Gene Family
JOURNAL: J. Clinical Investigation
VOLUME: 76
ISSUE: 12
PAGES: 2420-2424
DATE: December, 1985
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1-485
US-08-618-485B-1

Query Match 95.6%; Score 65; DB 4; Length 458;
Best Local Similarity 92.9%; Pred. No. 0.00056;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPTELAKLVNKRSE 14
DB 418 TPTELAKLVNKRSD 431

RESULT 4
5177002-1
PATENT No. 5177002
APPLICANT: YAMAMOTO, NOBUTO
TITLE OF INVENTION: IN VITRO ENZYMATIC CONVERSION OF
GLYCOSYLATED HUMAN VITAMIN D BINDING PROTEIN TO A POTENT
MACROPHAGE ACTIVATING FACTOR
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
FILING DATE: 31-AUG-1990
PRIOR APPLICATION NUMBER: 439,223
FILING DATE: 20-NOV-1989
SEQ ID NO: 1:
LENGTH: 458

Query Match 95.6%; Score 65; DB 6; Length 458;
Best Local Similarity 92.9%; Pred. No. 0.00056;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TPTELAKLVNKRSE 14
DB 418 TPTELAKLVNKRSD 431

RESULT 5
5177002-2
PATENT No. 5177002
APPLICANT: YAMAMOTO, NOBUTO
TITLE OF INVENTION: IN VITRO ENZYMATIC CONVERSION OF
GLYCOSYLATED HUMAN VITAMIN D BINDING PROTEIN TO A POTENT
MACROPHAGE ACTIVATING FACTOR
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 439,223
FILING DATE: 20-NOV-1989
SEQ ID NO: 2:
LENGTH: 458

Query Match 86.8%; Score 59; DB 6; Length 458;
Best Local Similarity 85.7%; Pred. No. 0.0071;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPTELAKLVNKRSE 14
DB 418 TPTELAKLVNKRSD 431

DB 418 TPTELAKLVNKRSD 431

RESULT 6
US-08-222-619-5
Sequence 5, Application US/08222619
Patent No. 5652352
GENERAL INFORMATION:
APPLICANT: Lichenstein, Henri
APPLICANT: Lyons, David
APPLICANT: Wirtel, Mark
APPLICANT: Wright, Samuel
TITLE OF INVENTION: Afamin: A Human Serum Albumin-like
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04075
FILING DATE:
CLASSIFICATION:

Query Match 86.8%; Score 59; DB 1; Length 474;
Best Local Similarity 85.7%; Pred. No. 0.0074;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPTELAKLVNKRSE 14
DB 434 TPTELAKLVNKRSD 447

RESULT 7
PCT-US95-04075-5
Sequence 5, Application PC/TUS9504075
GENERAL INFORMATION:
APPLICANT: AMGEN INC.
TITLE OF INVENTION: Afamin: A Human Serum Albumin-like
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Center, Patent Operations/RRC
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: U.S.
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04075
FILING DATE:
CLASSIFICATION:

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-04075-5

Query Match 86.8%; Score 59; DB 5; Length 474;
Best Local Similarity 85.7%; Pred. No. 0.0074; 1; Mismatches 1; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPTETAKLVNKRSE 14
DB 434 TPTETAKLVNKRSD 447

RESULT 8
US-09-585-858-37
Sequence 37, Application US/09585858
Patent No. 6492161
GENERAL INFORMATION:
APPLICANT: Sigridur Hjorleifsdottir
APPLICANT: Gudmundur O. Hreggvidsson
APPLICANT: Olafur H. Fridjonsson
APPLICANT: Arnthor Aevardsson
APPLICANT: Jakob K. Kristjansson
TITLE OF INVENTION: Bacteriophage RM376 of a Thermophilic
TITLE OF INVENTION: Host Organism
FILE REFERENCE: 2739.1001-001
CURRENT APPLICATION NUMBER: US/09/585,858
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: 60/137,120
PRIOR FILING DATE: 1999-06-02
NUMBER OF SEQ ID NOS: 73
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 37
LENGTH: 898
TYPE: PRT
ORGANISM: Bacteriophage RB69
US-09-585-858-37

Query Match 60.3%; Score 41; DB 4; Length 898;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTEIATLVNKRSE 14
DB 472 PTEITVFNKRKE 484

RESULT 9
US-08-311-731A-351
Sequence 351, Application US/08311731A
Patent No. 6583266
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 351:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Mycobacterium leprae
US-08-311-731A-351

Query Match 57.4%; Score 39; DB 4; Length 82;
Best Local Similarity 57.1%; Pred. No. 5.1;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 TPTETAKLVNKRSE 14
DB 40 TPTETAKLVNKRSE 53

RESULT 10
US-08-311-731A-76
Sequence 76, Application US/08311731A
Patent No. 6583266
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-3500
TELEFAX: 617/720-2441
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-76

Query Match 57.4%; Score 39; DB 4; Length 99;
Best Local Similarity 57.1%; Pred. No. 6.3;
Matches 8; Conservative 3; Mismatches 0; Gaps 0;

OY 1 TPTLAKLVNKRSE 14
Db 24 TPTLAKLVNKRVT 37

RESULT 11

US-08-311-731A-295
Sequence 295, Application US/08311731A
Patent No. 658326
GENERAL INFORMATION:
APPLICANT: SMITH, DOUGLAS
APPLICANT: MAO, JEN-I
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 411
CORRESPONDENCE ADDRESS:
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
STREET: 600 ATLANTIC AVENUE
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,731A
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: GATES, EDWARD R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: C0044/7125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/720-2441
TELEFAX: 617/720-3500
INFORMATION FOR SEQ ID NO: 295:
SEQUENCE CHARACTERISTICS:
LENGTH: 115 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULAR TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Mycobacterium leprae
US-08-311-731A-295

Query Match 57.4%; Score 39; DB 4; Length 115;
Best Local Similarity 57.1%; Pred. No. 7.4;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 TPTLAKLVNKRSE 14
Db 40 TPTLAKLVNKRVT 53

RESULT 12

US-08-949-637-2
Sequence 2, Application US/08949637
Patent No. 5910414
GENERAL INFORMATION:

APPLICANT: Gwynn, Michael
APPLICANT: Kallendar, Howard
TITLE OF INVENTION: No. 5910414el Topoisomerase I
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/949,637
FILING DATE: 14-OCT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/028,370
FILING DATE: 15-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmil, Edward R.
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50566
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-949-637-2

Query Match 57.4%; Score 39; DB 2; Length 496;
Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 PTEBLAKLVNK 11
Db 424 PTEBLGRIYVK 433

RESULT 13

US-09-291-488-2
Sequence 2, Application US/09291488
Patent No. 6251387
GENERAL INFORMATION:
APPLICANT: Gwynn, Michael
APPLICANT: Kallendar, Howard
TITLE OF INVENTION: No. 6251387el Topoisomerase I
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/291,488
FILING DATE:

```
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/949,637
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Glimm, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50566
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-291-488-2
```

```
Query Match      57.4% Score 39; DB 3; Length 496;
Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 PTEIAKLUNK 11
      |||||:||||
Db      424 PTEIGRIYVK 433
```

```
RESULT 14
US-09-153-277-4
Sequence 4, Application US/09153277
Patent No. 6331411
GENERAL INFORMATION:
APPLICANT: Kallender, Howard
APPLICANT: Gwynn, Michael
APPLICANT: Sylvester, Daniel
APPLICANT: Katz, Lisa
APPLICANT: Warren, Richard L.
APPLICANT: Traini, Christopher M.
TITLE OF INVENTION: TOPA
FILE REFERENCE: P50566-01
CURRENT APPLICATION NUMBER: US/09/153,277
CURRENT FILING DATE: 1998-09-15
EARLIER APPLICATION NUMBER: 08/949,637
EARLIER FILING DATE: 1997-10-14
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 507
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-153-277-4
```

```
Query Match      57.4% Score 39; DB 4; Length 507;
Best Local Similarity 70.0%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 PTEIAKLUNK 11
      |||||:||||
Db      423 PTEIGRIYVK 432
```

```
RESULT 15
US-09-153-277-2
Sequence 2, Application US/09153277
Patent No. 6331411
GENERAL INFORMATION:
APPLICANT: Kallender, Howard
APPLICANT: Gwynn, Michael
APPLICANT: Sylvester, Daniel
APPLICANT: Katz, Lisa
```

```
APPLICANT: Warren, Richard L.
APPLICANT: Traini, Christopher M.
TITLE OF INVENTION: TOPA
FILE REFERENCE: P50566-01
CURRENT APPLICATION NUMBER: US/09/153,277
CURRENT FILING DATE: 1998-09-15
EARLIER APPLICATION NUMBER: 08/949,637
EARLIER FILING DATE: 1997-10-14
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 695
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-153-277-2
```

```
Query Match      57.4% Score 39; DB 4; Length 695;
Best Local Similarity 70.0%; Pred. No. 53;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      2 PTEIAKLUNK 11
      |||||:||||
Db      505 PTEIGRIYVK 514
```

```
Search completed: May 21, 2004, 18:04:17
Job time : 25 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 17:59:12 ; Search time 55 Seconds
(without alignments)
71.921 Million cell updates/sec

Title: US-10-045-673a-1

Sequence: 1 TPEELAKLVKRSK 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	14	5	ABB79890 Vitamin D
2	65	95.6	89	2	AAW10562 Vitamin D
3	65	95.6	94	2	AAW10563 Vitamin D
4	65	95.6	458	2	AAW22278 Human gp.
5	65	95.6	458	2	AAW56975 Variant v
6	65	95.6	458	2	AAW10561 Vitamin D
7	60	88.2	474	6	ABP58017 Human ser
8	59	86.8	458	2	AAW22279 Human gp.
9	59	86.8	458	2	AAW56976 Variant v
10	44	64.7	314	2	AAW48469 Human hme
11	44	64.7	476	4	ABG15879 Novel hum
12	43	63.2	705	6	ABJ25543 Aspergill
13	43	63.2	748	6	ABJ26143 Aspergill
14	41	60.3	898	4	ABJ26143 Aspergill
15	40	58.8	1124	5	ABB91986 Herdiciida
16	39	57.4	61	3	ABAB16799 Bacteriop
17	39	57.4	82	7	ADB74602 Mycobacte
18	39	57.4	98	7	ADB74327 Mycobacte
19	39	57.4	115	7	ADB74546 Mycobacte
20	39	57.4	496	2	AAW55962 Streptoco
21	39	57.4	507	3	AAW70631 S. pneumo
22	39	57.4	695	3	AAW70630 S. pneumo
23	39	57.4	701	6	ABU46085 Protein e
24	39	57.4	705	6	ABU01695 S. pneumo
25	39	57.4	709	5	ABP26371 Streptoco

26	39	57.4	709	6	ABU46649 Protein e
27	39	57.4	710	5	ABBS4564 Lactococc
28	38	55.9	151	4	AAU09101 Novel hum
29	38	55.9	200	2	AAW97594 Cellulose
30	38	55.9	216	5	ABG69360 Babesia a
31	38	55.9	259	5	ABG69363 Babesia a
32	38	55.9	441	4	ABG28629 Novel hum
33	38	55.9	528	6	ABM68206 Phototrab
34	38	55.9	556	2	AAW06620 Sulfotolu
35	38	55.9	556	2	AAW2755 Trehalose
36	38	55.9	711	6	AAW45138 Protein e
37	38	55.9	729	6	ABU28842 Protein e
38	38	55.9	1579	5	ABW91088 Herdiciida
39	38	55.9	1750	6	ABG21521 Novel hum
40	38	55.9	1750	6	ABG00788 Polypept1
41	38	55.9	1784	4	ABG24772 Novel hum
42	38	55.9	1946	6	ABU35023 Protein e
43	37	54.4	65	3	ABP40132 Staphyloc
44	37	54.4	109	3	AAW57744 Arabidops
45	37	54.4	130	3	AAW57743 Arabidops

ALIGNMENTS

RESULT 1	ABB79890	standard, peptide, 14 AA.
ID	ABB79890	
XX	ABB79890	
AC	29-NOV-2002	(first entry)
XX		
DE	Vitamin D binding protein domain III peptide fADP.	
XX		
KM	Vitamin D binding protein; fADP, bone; osteopathic; antiarthritic;	
XX	vulnerary; osteoporosis; therapy.	
OS	Homo sapiens.	
XX		
FM	Key	Location/Qualifiers
FT	Modified-site	3 /note="glycosylated with N-acetylglucosamine"
XX		
XX	WC002058589-A2.	
PD	01-AUG-2002.	
XX		
PF	09-NOV-2001; 2001MO-US050471.	
XX		
PR	09-NOV-2000; 2000US-0247464P.	
XX		
PA	(UYNE-) UNIV NORTHEASTERN OHIO.	
XX		
PI	Schneider GB, Popoff SN, Safadi F;	
XX	WPI, 2002-666885/71.	
DR		
XX		
PT	New peptide used for increasing bone density e.g. in treatment of	
XX	osteoporosis comprises specified amino acid sequence.	
PS	Claim 1, Fig 1; 49pp; English.	
XX		
CC	The present peptide, designated fADP, is derived from domain III of a	
CC	human vitamin D binding protein. It comprises a sugar residue, especially	
CC	an N-acetylglucosamine, attached at the Thr-3 residue, or lacks a sugar	
CC	molety. A claimed peptide for increasing bone density in a mammal	
CC	comprises the first 3-13 amino acids of fADP, or has at least 70%	
CC	(preferably at least 90%) homology to fADP. The peptide is used in a	
CC	claimed method for promoting bone deposition. This is useful in the	
CC	treatment of disorders involving bone loss, particularly osteoporosis,	
CC	osteogenesis imperfecta, osteopenia, bone fracture, reconstruction of	
CC	bone after tumour removal to achieve spine and other joint fusion, bone	

CC loss due to vascular necrosis, bone necrosis, in the repair of
 CC congenital, trauma induced, and oncologic resection induced defects, in
 CC cosmetic plastic surgery, in bone-involved wound healing and related
 CC repair, in the treatment of periodontal disease and in other tooth repair
 CC processes, and in the prevention and treatment of osteoarthritis. The
 CC peptide provides regrowth of host bone and an environment to attract bone
 CC -forming cells, stimulate growth of bone-forming cells, induce
 CC differentiation of progenitor bone-forming cells and supports the
 CC regeneration of the periodontal ligament and attachment apparatus that
 CC connects bone and teeth

XX SQ Sequence 14 AA;

Query Match 100.0%; Score 68; DB 5; Length 14;
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPTTELAKLVNKRSE 14
 |||||
 DB 1 TPTTELAKLVNKRSE 14

RESULT 2

ID AAW10562 standard; protein; 89 AA.

XX AAW10562;

DT 24-SEP-1997 (first entry)

XX Vitamin D3-binding protein domain III CDMAF1 from cloned Gc protein.

XX Cloned Gc protein-macrophage activating factor; human; baculovirus;
 KM cancer; viral infection; osteoporosis; immunisation; vaccination; HIV;
 KM Epstein-Barr virus; herpes zoster infection; alpha-Ag.

XX Homo sapiens.

XX MO9640903-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008867.

XX 07-JUN-1995; 95US-00478121.

XX 19-MAR-1996; 96US-00618485.

XX (YAMA/) YAMAMOTO N.

XX Yamamoto N;

XX WPI; 1997-087061/08.

XX Cloning vitamin D3 binding protein or its domain III in baculovirus - to
 PT produce macrophage activating factors, useful in cancer, viral infection
 PT and osteoporosis treatment.

XX Claim 19; Fig 4; 43pp; English.

XX The present sequence represents the cloned domain III-derived macrophage
 CC activating factor (CDMAF1), which is from a vitamin D3 binding protein
 CC (Gcp). Full length cDNA encoding the human Gc protein, isolated from a
 CC human liver cDNA library in bacteriophage lambda, was cloned into a
 CC baculovirus expression system to produce the cloned protein
 CC recombinantly. GcMAFc, CDMAF and GcMAF (similar to GcMAFc, but prepared
 CC from serum rather than recombinant protein) are used to treat cancer, or
 CC HIV, Epstein-Barr virus or herpes zoster infection, and to promote bone
 CC formation in osteoporosis. They are also adjuvants for immunisation, or
 CC vaccination. Alpha-Ag is present in all patients with cancer and HIV
 CC infection, and its level is inversely correlated with precursor activity
 CC of plasma GcP, so measurement of alpha-Ag may be useful for diagnosis and
 CC prognosis

SQ Sequence 89 AA;

Query Match 95.6%; Score 65; DB 2; Length 89;
 Best Local Similarity 92.9%; Pred. No. 0.0002;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPTTELAKLVNKRSE 14
 |||||
 DB 49 TPTTELAKLVNKRSD 62

RESULT 3

ID AAW10563 standard; protein; 94 AA.

XX AAW10563;

DT 24-SEP-1997 (first entry)

XX Vitamin D3-binding protein domain III CDMAF2 from cloned Gc protein.

XX Cloned Gc protein-macrophage activating factor; human; baculovirus;
 KM cancer; viral infection; osteoporosis; immunisation; vaccination; HIV;
 KM Epstein-Barr virus; herpes zoster infection; alpha-Ag.

XX Homo sapiens.

XX MO9640903-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US008867.

XX 07-JUN-1995; 95US-00478121.

XX 19-MAR-1996; 96US-00618485.

XX (YAMA/) YAMAMOTO N.

XX Yamamoto N;

XX WPI; 1997-087061/08.

XX Cloning vitamin D3 binding protein or its domain III in baculovirus - to
 PT produce macrophage activating factors, useful in cancer, viral infection
 PT and osteoporosis treatment.

XX Claim 20; Fig 6; 43pp; English.

XX The present sequence represents the cloned domain III-derived macrophage
 CC activating factor (CDMAF2), which is from a vitamin D3 binding protein
 CC (Gcp). Full length cDNA encoding the human Gc protein, isolated from a
 CC human liver cDNA library in bacteriophage lambda, was cloned into a
 CC baculovirus expression system to produce the cloned protein
 CC recombinantly. GcMAFc, CDMAF and GcMAF (similar to GcMAFc, but prepared
 CC from serum rather than recombinant protein) are used to treat cancer, or
 CC HIV, Epstein-Barr virus or herpes zoster infection, and to promote bone
 CC formation in osteoporosis. They are also adjuvants for immunisation, or
 CC vaccination. Alpha-Ag is present in all patients with cancer and HIV
 CC infection, and its level is inversely correlated with precursor activity
 CC of plasma GcP, so measurement of alpha-Ag may be useful for diagnosis and
 CC prognosis

XX SQ Sequence 94 AA;

Query Match 95.6%; Score 65; DB 2; Length 94;
 Best Local Similarity 92.9%; Pred. No. 0.00022;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPTTELAKLVNKRSE 14
 |||||
 DB 54 TPTTELAKLVNKRSD 67

RESULT 4

AAR22278 standard; protein; 458 AA.

AAR22278;

27-JUL-1992 (first entry)

Human gp.-specific component phenotype Gc1.

Macrophage activating factor; MAF; galactosidase; sialidase; mannosidase;

B cells; T cells; cancer; immunodeficient diseases;

vitamin D binding protein.

Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 152 /note= "Gly in Gc2"

FT Misc-difference 311 /note= "Glu in Gc2"

FT Misc-difference 416 /note= "Asp in Gc2"

FT Misc-difference 420 /note= "Lys in Gc2"

W09204459-A.

19-MAR-1992.

29-AUG-1991; 91WO-US006172.

31-AUG-1990; 90US-00576248.

(YAMA/) YAMAMOTO N.

Yamamoto N;

WPI; 1992-114364/14.

Prepn. of new macrophage activating factor - by contacting human gp.-

specific component with glycosidase(s), useful for treating immuno-

deficient diseases and cancer.

Disclosure; Fig 1; 29pp; English.

The vitamin D binding protein Gc1 is an evolutionary conserved

glycoprotein. Polymorphism of the Gc protein was demonstrated by gel

electrophoretic analysis, which revealed two major phenotypes: Gc1 and

Gc2 (AAR22279). The Gc protein may be purified by a variety of means from

blood, e.g. by 25-hydroxy vitamin D3-Sepharose affinity chromatography or

actin-agarose affinity chromatography. Gc1 may be converted to MAF by the

action of glycosidases of B and T cells, e.g. by contacting Gc1 in vitro

with beta-galactosidase or beta-gal in combination with sialidase and/or

alpha-mannosidase. The MAF may be produced in large amounts with high

potency. MAF may be used for inducing macrophage activation, partic. in

patients with immunodeficient diseases, cancer and other diseases

characterised by impaired B- or T-cell function

Sequence 458 AA;

Query Match 95.6%; Score 65; DB 2; Length 458;

Best Local Similarity 92.9%; Pred. No. 0.0013;

Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 TPTETLAKLVNKRSE 14

418 TPTETLAKLVNKRSD 431

RESULT 5

AAR56975

ID AAR56975 standard; protein; 458 AA.

AAR56975;

25-MAR-2003 (revised)

16-FEB-1995 (first entry)

Variant vitamin D binding protein (DBPge/gm).

Vitamin D binding protein; DBP; macrophage activating factor;

oligosaccharide; galactose; alpha mannose; sialic acid;

beta-galactosidase; alpha-mannosidase; sialidase; MAF.

Homo sapiens.

US5326749-A.

05-JUL-1994.

04-JAN-1993; 93US-00000320.

20-NOV-1989; 89US-00439223.

31-AUG-1990; 90US-00576248.

30-SEP-1991; 91US-00767742.

(YAMA/) YAMAMOTO N.

Yamamoto N;

WPI; 1994-217073/26.

Macrophage activating factor - prepared by treating glycosated vitamin D-

binding protein with glycoside(s).

Disclosure; Fig 1; 12pp; English.

Vitamin D binding protein (DBP) is converted to a macrophage activating

factor by the action of B and T cell glycosidases. The polymorphic DBP

phenotypes are expressed inter alia as differences in the oligosaccharide

attached to the polypeptide portion of the DBP molecule. All three

principal DBP types - DBPgm, DBPgs and DBPg differ in the nature of the

appended oligosaccharide. DBPgs which possesses an oligosaccharide moiety

which includes galactose and sialic acid residues can be converted to the

macrophage activating factor by contact with beta-galactosidase and

sialidase. DBPgm which is believed to possess an oligosaccharide moiety

which includes galactose and alpha-mannose residues is contacted with

beta-galactosidase and alpha-mannosidase to produce the macrophage

activating factor and DBPg which is believed to possess an

oligosaccharide moiety which includes galactose without sialic acid or

alpha-mannose is contacted with beta-galactosidase alone to form the

macrophage activating factor. The resulting factor is of high potency.

(Updated on 25-MAR-2003 to correct PF field.)

Sequence 458 AA;

Query Match 95.6%; Score 65; DB 2; Length 458;

Best Local Similarity 92.9%; Pred. No. 0.0013;

Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 TPTETLAKLVNKRSE 14

418 TPTETLAKLVNKRSD 431

RESULT 6

AAM10561

ID AAM10561 standard; protein; 458 AA.

AAM10561;

24-SEP-1997 (first entry)

Vitamin D3-binding protein GcWAFc from cloned Gc protein.

KM Cloned Gc protein-macrophage activating factor; human; baculovirus;
 KW cancer; viral infection; osteoporosis; immunisation; vaccination; HIV;
 KM Epstein-Barr virus; herpes zoster infection; alpha-Ag.
 XX Homo sapiens.
 OS
 XX WO640903-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 05-JUN-1996; 96WO-US008867.
 XX
 PR 07-JUN-1995; 95US-00478121.
 PR 19-MAR-1996; 96US-00618485.
 XX
 PA (YAMA/) YAMAMOTO N.
 XX
 PI Yamamoto N;
 DR WPI; 1997-087061/08.
 XX
 PT Cloning vitamin D3 binding protein or its domain III in baculovirus - to
 PT produce macrophage activating factors, useful in cancer, viral infection
 PT and osteoporosis treatment.
 XX
 PS Claim 18; Fig 2; 43pp; English.
 XX
 CC The present sequence represents the cloned Gc protein-derived macrophage
 CC activating factor (GcMAFc), which is a vitamin D3 binding protein (GcP).
 CC Full length cDNA encoding the human Gc protein, isolated from a human
 CC liver cDNA library in bacteriophage lambda, was cloned into a baculovirus
 CC expression system to produce the cloned protein recombinantly. GcMAFc,
 CC CdMAr and GcMAr (similar to GcMAFc, but prepared from serum rather than
 CC recombinant protein) are used to treat cancer, or HIV, Epstein-Barr virus
 CC or herpes zoster infection, and to promote bone formation in
 CC osteoporosis. They are also adjuvants for immunisation, or vaccination.
 CC Alpha-Ag is present in all patients with cancer and HIV infection, and
 CC its level is inversely correlated with precursor activity of Plasma GcP,
 CC so measurement of alpha-Ag may be useful for diagnosis and prognosis
 XX
 SQ Sequence 458 AA;
 Query Match 95.6%; Score 65; DB 2; Length 458;
 Best Local Similarity 92.9%; Pred. No. 0.0013;
 Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TPTELAKLVNKRSE 14
 Db 418 TPTELAKLVNKRSD 431
 RESULT 7
 ABP58017
 ID ABP58017 standard; protein; 474 AA.
 XX
 AC ABP58017;
 XX
 DT 11-FEB-2003 (first entry)
 XX
 DE Human serum vitamin D binding protein.
 XX
 KM Prostate cancer; marker; vitamin D binding protein; VDBP; human;
 KM diagnosis; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Peptide Location/Qualifiers
 FT 50..65
 FT /note="marker peptide"
 FT 51..65
 FT /note="marker peptide"
 FT 88..94
 FT /note="marker peptide"

FT Peptide 115..119
 FT /note="marker peptide"
 FT 208..218
 FT /note="marker peptide"
 FT 285..292
 FT /note="marker peptide"
 FT 342..352
 FT /note="marker peptide"
 FT 346..352
 FT /note="marker peptide"
 FT 353..363
 FT /note="marker peptide"
 FT 354..363
 FT /note="marker peptide"
 FT 365..370
 FT /note="marker peptide"
 FT 389..393
 FT /note="marker peptide"
 XX
 XX WO200275314-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 30-NOV-2001; 2001WO-US045031.
 XX
 PR 30-NOV-2000; 2000US-0250284P.
 PR 08-NOV-2001; 2001US-0344948P.
 XX
 PA (MATR-) MATRITTECH INC.
 XX
 PI Hlavaty J, Briggsman JV;
 DR WPI; 2003-067369/06.
 DR N-PsDB; ABV76007.
 XX
 PT Diagnosing or treating prostate cancer by detecting in a sample isolated
 PT from the individual the presence of prostate cancer-associated protein.
 XX
 PS Disclosure; Fig 2; 63pp; English.
 XX
 CC The present sequence is the protein sequence of an allele of human human
 CC serum vitamin D binding protein (VDBP). The invention provides a novel
 CC human protein that includes amino acid sequences (see ABP58005-16) that
 CC are also found in VDBP. These marker sequences can be used to distinguish
 CC VDBP-related proteins from other proteins. VDBP-related proteins are
 CC detectable at a higher concentration in serum from a mammal, e.g. a
 CC human, with prostate cancer relative to serum from a healthy mammal and
 CC can therefore be used as prostate cancer markers. They permit the rapid
 CC detection, preferably before metastases occur, of prostate cancer. A
 CC target prostate cancer-associated protein may be detected using a
 CC labelled antibody capable of binding specifically to the protein.
 CC Prostate cancer-associated proteins, and nucleic acids encoding them, are
 CC also useful as targets for treating prostate cancer, and as indicators
 CC for monitoring the efficiency of prostate cancer therapy
 XX
 SQ Sequence 474 AA;
 Query Match 88.2%; Score 60; DB 6; Length 474;
 Best Local Similarity 85.7%; Pred. No. 0.012;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TPTELAKLVNKRSE 14
 Db 434 TPTELAKLVNKRSD 447
 RESULT 8
 AAR22279
 ID AAR22279 standard; protein; 458 AA.
 XX
 AC AAR22279;
 XX
 DT 27-JUL-1992 (first entry)

PD 23-SEP-1999.
 XX
 PF 20-MAR-1998; 98DE-01013835.
 XX
 PR 20-MAR-1998; 98DE-01013835.
 XX
 PA (META-) METAGEN GBS GENOMFORSCHUNG MBH.
 PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 XX WPI; 1999-528979/45.
 DR N-PSDB; AA233550.
 XX
 PT Human nucleic acid sequences and protein products from normal breast
 PT tissue, useful for breast cancer therapy.
 XX
 PS Claim 28; 165; 206pp; German.
 CC This invention describes novel human nucleic acid sequences from normal
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer and for treating illnesses associated with
 CC fat metabolism. AA48456-Y48539 represent protein fragments encoded by
 CC the expressed sequence tags described in the method of the invention
 CC
 SQ Sequence 314 AA;
 Query Match 64.7%; Score 44; DB 2; Length 314;
 Best Local Similarity 69.2%; Pred. No. 8.1;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 PTEIAKLVNKRSE 14
 Db 163 PSEIRNLVNRKSE 175
 :||| |||||
 RESULT 11
 ABG15879
 ID ABG15879 standard; protein; 476 AA.
 XX
 AC ABG15879;
 XX
 DT 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #15870.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 21-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS80066.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess

PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 46238; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a tool
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensic, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WPIO at
 CC ftp.wpi.int/pub/published_poc_sequences
 CC
 SQ Sequence 476 AA;
 Query Match 64.7%; Score 44; DB 4; Length 476;
 Best Local Similarity 69.2%; Pred. No. 13;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 PTEIAKLVNKRSE 14
 Db 325 PSEIRNLVNRKSE 337
 :||| |||||
 RESULT 12
 ABJ25543
 ID ABJ25543 standard; protein; 705 AA.
 XX
 AC ABJ25543;
 XX
 DT 16-APR-2003 (first entry)
 DE Aspergillus fumigatus essential gene protein #201.
 XX
 KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
 KW cancer; contamination; biofilm; antibody; immune response.
 XX
 OS Aspergillus fumigatus.
 XX
 PN W0200286090-A2.
 PD 31-OCT-2002.
 XX
 PF 23-APR-2002; 2002WO-US013142.
 XX
 PR 23-APR-2001; 2001US-0285697P.
 PR 27-APR-2001; 2001US-0287066P.
 PR 05-JUN-2001; 2001US-0295890P.
 PR 09-JUL-2001; 2001US-0303899P.
 PR 31-AUG-2001; 2001US-0316362P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
 XX WPI; 2003-093124/08.
 DR
 XX
 PT New purified or isolated nucleic acids of essential genes of Aspergillus
 PT fumigatus, useful for treating or preventing infections by A. fumigatus,
 PT or for treating a non-infectious disease in a subject e.g. cancer.

XX PS Disclosure; Page: 175pp; English.

CC The invention relates to novel purified or isolated nucleic acids of

CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of

CC the invention are used to treat or prevent infections by a pathogenic

CC organism such as *A. fumigatus*, to treat a non-infectious disease in a

CC subject (e.g. cancer), to prevent or contain contamination of an object

CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a

CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for

CC expressing recombinant protein for characterisation, screening or

CC therapeutic use, as markers for host tissues in which the pathogenic

CC organisms invade or reside, for comparing with the DNA sequence of *A.*

CC *fumigatus* to identify duplicated genes or paralogues having the same or

CC similar biochemical activity and/or function, for comparing with DNA

CC sequences of other related or distant pathogenic organisms to identify

CC potential orthologous essential or virulence genes, for selecting and

CC making oligomers for attachment to a nucleic acid array for examination

CC of expression patterns, for raising anti-protein antibodies, as an

CC antigen to raise anti-DNA antibodies or to elicit another immune

CC response, and for identifying polynucleotides encoding the other protein

CC with which binding occurs or to identify inhibitors of the binding

CC interaction. The polypeptides may be used to raise antibodies or to

CC elicit immune response, as a reagent in assays designed to quantitatively

CC determine levels of the protein in biological fluids, as a marker for

CC host tissues in which pathogenic organism invade or reside, and to

CC isolate correlative receptors or ligands in the case of virulence

CC factors. This sequence represents a protein of one of the essential genes

CC of *Aspergillus fumigatus* of the invention

XX SQ Sequence 705 AA;

Query Match 63.2%; Score 43; DB 6; Length 705;

Best Local Similarity 64.3%; Pred. No. 31;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 TPTETAKLVNKRSE 14

Db 562 TPADLAKLAEIRSE 575

RESULT 13

ABJ26143

ID ABJ26143 standard; protein; 748 AA.

XX AC ABJ26143;

XX DT 16-APR-2003 (first entry)

XX DE *Aspergillus fumigatus* essential gene protein #801.

XX KW Fungicide; cytostatic; essential gene; *Aspergillus fumigatus*; infection;

XX KM cancer; contamination; biofilm; antibody; immune response.

XX OS *Aspergillus fumigatus*.

XX PN WO200286090-A2.

XX PD 31-OCT-2002.

XX PF 23-APR-2002; 2002WO-US013142.

XX PR 23-APR-2001; 2001US-0285697P.

XX PR 27-APR-2001; 2001US-0287066P.

XX PR 05-JUN-2001; 2001US-0295890P.

XX PR 09-JUL-2001; 2001US-0303899P.

XX PR 31-AUG-2001; 2001US-0316362P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PJ Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;

XX WPI, 2003-093124/08.

XX PT New purified or isolated nucleic acids of essential genes of *Aspergillus*

XX *fumigatus*, useful for treating or preventing infections by *A. fumigatus*,

XX or for treating a non-infectious disease in a subject e.g. cancer.

XX PS Disclosure; Page: 175pp; English.

CC The invention relates to novel purified or isolated nucleic acids of

CC essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of

CC the invention are used to treat or prevent infections by a pathogenic

CC organism such as *A. fumigatus*, to treat a non-infectious disease in a

CC subject (e.g. cancer), to prevent or contain contamination of an object

CC by *A. fumigatus*, or to prevent or inhibit formation on a surface of a

CC biofilm comprising *A. fumigatus*. The polynucleotides are useful for

CC expressing recombinant protein for characterisation, screening or

CC therapeutic use, as markers for host tissues in which the pathogenic

CC organisms invade or reside, for comparing with the DNA sequence of *A.*

CC *fumigatus* to identify duplicated genes or paralogues having the same or

CC similar biochemical activity and/or function, for comparing with DNA

CC sequences of other related or distant pathogenic organisms to identify

CC potential orthologous essential or virulence genes, for selecting and

CC making oligomers for attachment to a nucleic acid array for examination

CC of expression patterns, for raising anti-protein antibodies, as an

CC antigen to raise anti-DNA antibodies or to elicit another immune

CC response, and for identifying polynucleotides encoding the other protein

CC with which binding occurs or to identify inhibitors of the binding

CC interaction. The polypeptides may be used to raise antibodies or to

CC elicit immune response, as a reagent in assays designed to quantitatively

CC determine levels of the protein in biological fluids, as a marker for

CC host tissues in which pathogenic organism invade or reside, and to

CC isolate correlative receptors or ligands in the case of virulence

CC factors. This sequence represents a protein of one of the essential genes

CC of *Aspergillus fumigatus* of the invention

XX SQ Sequence 748 AA;

Query Match 63.2%; Score 43; DB 6; Length 748;

Best Local Similarity 64.3%; Pred. No. 33;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 TPTETAKLVNKRSE 14

Db 605 TPADLAKLAEIRSE 618

RESULT 14

AAB46728

ID AAB46728 standard; protein; 898 AA.

XX AC AAB46728;

XX DT 11-SEP-2003 (revised)

XX DT 12-APR-2001 (first entry)

XX DE Bacteriophage RB69 DNA polymerase protein fragment SEQ ID NO 37.

XX KM Genome; thermophilic enzyme; washing powder; bleaching.

XX OS Enterobacteria phage RB69.

XX PN WO200075335-A2.

XX PD 14-DEC-2000.

XX PF 02-JUN-2000; 2000WO-IB000893.

XX PR 02-JUN-1999; 99US-0137120P.

XX PA (DECO-) DECODE GENETICS EHF.

XX PJ Hjortleifsdottir S, Hreggvidsson GO, Fridjonsson OH, Aevartson A;

XX Kristjansson JK;

DR WPI; 2001-061727/07.

XX Isolated nucleic acid from bacteriophage RM 378 encoding enzymes useful
PT in recombinant DNA technology.

XX Disclosure; Fig 3A-P; 42pp; English.

CC This invention describes a novel isolated nucleic molecule (I) comprising
CC the genome of bacteriophage RM 378. The invention also describes (1) an
CC isolated nucleic acid which encodes a polypeptide obtainable from
CC bacteriophage RM 378, or its active derivative or fragment; (2) an
CC isolated bacteriophage RM 378 (II); (3) a DNA construct (III) comprising
CC operatively linked to a regulatory sequence; (4) a host cell comprising
CC (III); and (5) an isolated polypeptide (IV) obtainable from (II), its
CC active derivative or fragment. Bacteriophage RM 378 is useful for
CC producing thermophilic enzymes useful in DNA research and commercial
CC settings (e.g. proteases and lipases used in washing powder, hydrolytic
CC enzymes used in bleaching). The isolated nucleic acid molecules and
CC vectors are useful in the manufacture of encoded polypeptide, as probes
CC for isolating homologous sequences (e.g. from other bacteriophage
CC species), as well as for detecting the presence of the bacteriophage in a
CC culture of host cells. The polypeptides can be used as a molecular weight
CC marker on SDS-PAGE gels or on molecular sieve gel filtration columns.
CC Because the host organism of the RM378 bacteriophage is a thermophile,
CC the enzymes and proteins of the RM378 bacteriophage are significantly
CC more thermostable than those of other (e.g. mesophilic) bacteriophages,
CC such as the T4 bacteriophage of *Escherichia coli*. The enhanced stability
CC of the enzymes and proteins of RM378 bacteriophage allows their use under
CC temperature conditions which would be prohibitive for other enzymes, thus
CC increasing the range of conditions which can be employed not only in DNA
CC research but also in commercial settings. (Updated on 11-SEP-2003 to
CC standardise OS field)

XX SQ Sequence 898 AA;

Query Match

Best Local Similarity 60.3%; Score 41; DB 4; Length 898;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 PTELAKLVNKRSE 14

DB 472 PTEITKVFVNORKE 484

RESULT 15

ABB91986

ID ABB91986 standard; protein; 1124 AA.

XX ABB91986;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1197.

XX Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP009892.

PR 28-AUG-2001; 2001WO-EP009892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidner M;

XX WPI; 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,

PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
XX organisms.

PS Claim 5; SEQ ID NO 1197; 261pp + Sequence listing; English.

CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides

XX SQ Sequence 1124 AA;

Query Match

Best Local Similarity 58.8%; Score 40; DB 5; Length 1124;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PTELAKLVN 10

DB 486 PTELAKLVN 494

Search completed: May 21, 2004, 18:01:52

Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 18:01:58 ; Search time 42 Seconds
(without alignments)
92.974 Million cell updates/sec

Title: US-10-045-673a-1
Sequence: 1 TPTELAKLVNKRSE 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1149313 seqs, 278921704 residues

Total number of hits satisfying chosen parameters: 1149313

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	14	US-10-045-673a-1	Sequence 1, Appl
2	65	95.6	89	US-09-826-463-2	GENERAL INFORMA
3	65	95.6	94	US-09-826-463-3	GENERAL INFORMA
4	65	95.6	458	US-09-826-463-1	GENERAL INFORMA
5	60	88.2	474	US-09-998-909-14	Sequence 14, Appl
6	44	64.7	58	US-10-029-386-34112	Sequence 34112, A
7	43	63.2	705	US-10-128-714-3201	Sequence 3201, Ap
8	43	63.2	748	US-10-128-714-8201	Sequence 8201, Ap
9	42	61.8	406	US-10-369-493-20429	Sequence 20429, A
10	41	60.3	856	US-10-401-403-196	Sequence 196, App
11	41	60.3	898	US-10-270-878-37	Sequence 37, Appl
12	41	60.3	898	US-10-270-878-37	Sequence 37, Appl
13	41	60.3	898	US-10-270-786-37	Sequence 37, Appl
14	41	60.3	898	US-10-270-710-37	Sequence 37, Appl
15	41	60.3	898	US-10-270-859-37	Sequence 37, Appl

16	41	60.3	898	US-10-270-846-37	Sequence 37, Appl
17	39	57.4	61	US-10-097-111-397	Sequence 397, App
18	39	57.4	701	US-10-282-122a-74009	Sequence 74009, A
19	39	57.4	709	US-10-282-122a-74573	Sequence 74573, A
20	39	57.4	710	US-10-369-493-18415	Sequence 18415, A
21	39	57.4	865	US-10-425-114-55769	Sequence 55769, A
22	39	57.4	1029	US-10-424-599-245114	Sequence 245114, A
23	39	57.4	1161	US-10-369-493-3660	Sequence 3660, Ap
24	38	55.9	151	US-09-969-384-22	Sequence 22, Appl
25	38	55.9	216	US-10-234-432-28	Sequence 28, Appl
26	38	55.9	259	US-10-234-432-31	Sequence 31, Appl
27	38	55.9	327	US-10-369-493-3029	Sequence 3029, Ap
28	38	55.9	711	US-10-282-122a-73062	Sequence 73062, A
29	38	55.9	729	US-10-282-122a-56766	Sequence 56766, A
30	38	55.9	729	US-10-369-493-23632	Sequence 23632, A
31	38	55.9	1750	US-10-243-552-920	Sequence 920, App
32	38	55.9	1946	US-10-282-122a-62947	Sequence 62947, A
33	37	54.4	493	US-10-425-114-57893	Sequence 57893, A
34	37	54.4	506	US-10-108-260a-4387	Sequence 4387, Ap
35	37	54.4	588	US-10-213-990-6	Sequence 6, Appl
36	37	54.4	692	US-10-282-122a-60918	Sequence 60918, A
37	37	54.4	1230	US-10-093-524-6	Sequence 6, Appl
38	37	54.4	1820	US-10-369-493-5274	Sequence 5274, Appl
39	37	54.4	1820	US-10-369-493-5900	Sequence 5900, Ap
40	36	52.9	51	US-10-424-599-164190	Sequence 164190, A
41	36	52.9	62	US-09-864-408a-6676	Sequence 6676, Ap
42	36	52.9	79	US-10-424-599-192519	Sequence 192519, A
43	36	52.9	178	US-10-282-122a-71863	Sequence 71863, A
44	36	52.9	212	US-10-424-599-230315	Sequence 230315, A
45	36	52.9	250	US-10-289-762-859	Sequence 859, App

ALIGNMENTS

RESULT 1
US-10-045-673a-1
; Sequence 1, Application US/10045673a
; Publication No. US20030229014A1
; GENERAL INFORMATION:
; APPLICANT: Schneider, Gary B
; APPLICANT: Popoff, Steven N
; APPLICANT: Safadi, Payez
; TITLE OF INVENTION: Agents and Methods for Promoting Bone Growth
; FILE REFERENCE: 25080/0400
; CURRENT APPLICATION NUMBER: US/10/045, 673a
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/247,464
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-045-673a-1

Query Match 100.0%; Score 68; DB 15; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPTELAKLVNKRSE 14
DB 1 TPTELAKLVNKRSE 14

RESULT 2
US-09-826-463-2
; GENERAL INFORMATION:
; APPLICANT: NOBUO YAMAMOTO
; TITLE OF INVENTION: PREPARATION OF POTENT
MACROPHAGE ACTIVATING FACTORS
DERIVED FROM CLONED VITAMIN D

BINDING PROTEIN AND ITS DOMAIN

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSSEE: CAESAR, RIVISE, BERNSTEIN,
COHEN & POKOTILOV, LTD.
STREET: 1635 Market Street, 12th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2212

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44 MB
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT VERSION 4.2

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/826,463
FILING DATE: 05-Apr-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/618,485B
FILING DATE: March 19, 1996
APPLICATION NUMBER: US 08/478,121
FILING DATE: 07-JUNE-1995

ATTORNEY/AGENT INFORMATION:
NAME: Robert S. Silver
REGISTRATION NUMBER: 35,681
REFERENCE/DOCKET NUMBER: Y1004/20002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2010
TELEFAX: (215) 751-1142
JOURNAL: J. Clinical Investigation
VOLUME: 76
ISSUE: 12
PAGES: 2420-2424
DATE: December, 1985

RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 4 and 5 TO 89
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-826-463-2

Query Match 95.6%; Score 65; DB 9; Length 89;
Best Local Similarity 92.9%; Pred. No. 0.00041;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPTELAKLVNKRSE 14
Db 49 TPTELAKLVNKRSD 62

RESULT 3
US-09-826-463-3

GENERAL INFORMATION:
APPLICANT: NOBUTO YAMAMOTO
TITLE OF INVENTION: PREPARATION OF POTENT
MACROPHAGE ACTIVATING FACTORS
DERIVED FROM CLONED VITAMIN D
BINDING PROTEIN AND ITS DOMAIN

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSSEE: CAESAR, RIVISE, BERNSTEIN,
COHEN & POKOTILOV, LTD.
STREET: 1635 Market Street, 12th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2212

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44 MB
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT VERSION 4.2

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/826,463

FILING DATE: 05-Apr-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/618,485B
FILING DATE: March 19, 1996
APPLICATION NUMBER: US 08/478,121
FILING DATE: 07-JUNE-1995

ATTORNEY/AGENT INFORMATION:
NAME: Robert S. Silver
REGISTRATION NUMBER: 35,681
REFERENCE/DOCKET NUMBER: Y1004/20002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2010
TELEFAX: (215) 751-1142
JOURNAL: J. Clinical Investigation
VOLUME: 76
ISSUE: 12
PAGES: 2420-2424
DATE: December, 1985

RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 10 TO 94
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-826-463-3

Query Match 95.6%; Score 65; DB 9; Length 94;
Best Local Similarity 92.9%; Pred. No. 0.00043;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPTELAKLVNKRSE 14
Db 54 TPTELAKLVNKRSD 67

RESULT 4
US-09-826-463-1

GENERAL INFORMATION:
APPLICANT: NOBUTO YAMAMOTO
TITLE OF INVENTION: PREPARATION OF POTENT
MACROPHAGE ACTIVATING FACTORS
DERIVED FROM CLONED VITAMIN D
BINDING PROTEIN AND ITS DOMAIN

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSSEE: CAESAR, RIVISE, BERNSTEIN,
COHEN & POKOTILOV, LTD.
STREET: 1635 Market Street, 12th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2212

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette-3.5 inch, 1.44 MB
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT VERSION 4.2

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/826,463
FILING DATE: 05-Apr-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/618,485B
FILING DATE: March 19, 1996
APPLICATION NUMBER: US 08/478,121
FILING DATE: 07-JUNE-1995

ATTORNEY/AGENT INFORMATION:
NAME: Robert S. Silver
REGISTRATION NUMBER: 35,681
REFERENCE/DOCKET NUMBER: Y1004/20002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 567-2010
TELEFAX: (215) 751-1142
JOURNAL: J. Clinical Investigation
VOLUME: 76
ISSUE: 12

PAGES: 2420-2424
DATE: December, 1985
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1-485
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-826-463-1

Query Match 95.6%; Score 65; DB 9; Length 458;
Best Local Similarity 92.9%; Pred. No. 0.0025;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TPTETALVVKRSE 14
DB 418 TPTETALVVKRSD 431

RESULT 5
US-09-998-909-14
; Sequence 14, Application US/09998909
; Patent No. US2002016464A1
; GENERAL INFORMATION:
; APPLICANT: Hlaavay, John
; APPLICANT: Brigsmann, Joseph
; TITLE OF INVENTION: Detection and Treatment of Prostate Cancer
; FILE REFERENCE: MTP-027
; CURRENT APPLICATION NUMBER: US/09/998,909
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 60/250,284
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-998-909-14

Query Match 88.2%; Score 60; DB 9; Length 474;
Best Local Similarity 85.7%; Pred. No. 0.021;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TPTETALVVKRSE 14
DB 434 TPTETALVVKRSD 447

RESULT 6
US-10-029-386-34172
; Sequence 34172, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34172
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AJ009613.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: SWISSPROT HIT: Q04279, EVALUATE 2.10e+00
US-10-029-386-34172

Query Match 64.7%; Score 44; DB 14; Length 58;
Best Local Similarity 69.2%; Pred. No. 1.7;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 TPTETALVVKRSE 14
DB 18 PSELRLVVKRSE 30

RESULT 7
US-10-128-714-3201
; Sequence 3201, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; FILE REFERENCE: Methods of Use
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3201
; LENGTH: 705
; TYPE: PRT
; ORGANISM: *Aspergillus fumigatus*
US-10-128-714-3201

Query Match 63.2%; Score 43; DB 14; Length 705;
Best Local Similarity 64.3%; Pred. No. 41;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TPTETALVVKRSE 14
DB 562 TPTETALVVKRSD 575

RESULT 8
US-10-128-714-8201
; Sequence 8201, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in *Aspergillus fumigatus* and
; FILE REFERENCE: Methods of Use
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899

```

; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8201
; LENGTH: 748
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8201

```

```

Query Match      63.2%; Score 43; DB 14; Length 748;
Best Local Similarity 64.3%; Pred. No. 44;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1 TPTSLAKLVNKRSE 14
      |||:|:|:|:|
DB      605 TPADLAKLAELRSE 618

```

```

RESULT 9
US-10-369-493-20429
; Sequence 20429, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianning
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20429
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Rhodospseudomonas palustris
US-10-369-493-20429

```

```

Query Match      61.8%; Score 42; DB 15; Length 406;
Best Local Similarity 81.8%; Pred. No. 34;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      3 TELAKLVNKRK 13
      |||:|:|:|:|
DB      45 TELAKLVNKRK 55

```

```

RESULT 10
US-10-401-403-196
; Sequence 196, Application US/10401403
; Publication No. US20040005599A1
; GENERAL INFORMATION:
; APPLICANT: Schoenbrunner, Nancy
; APPLICANT: Myers, Thomas
; APPLICANT: Gelfand, David
; TITLE OF INVENTION: THERMOSTABLE OR THERMOACTIVE DNA POLYMERASE MOLECULES
; FILE REFERENCE: 21314-US1
; CURRENT APPLICATION NUMBER: US/10/401,403
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: US 60/369,815
; PRIOR FILING DATE: 2002-04-02
; NUMBER OF SEQ ID NOS: 203
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 196
; LENGTH: 856
; TYPE: PRT

```

```

; ORGANISM: Bacteriophage RB69
US-10-401-403-196

```

```

Query Match      60.3%; Score 41; DB 15; Length 856;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 PTEILKLVNKRSE 14
      |||:|:|:|:|
DB      472 PTEITKVFQNRKE 484

```

```

RESULT 11
US-10-270-875-37
; Sequence 37, Application US/10270875
; Publication No. US20030082741A1
; GENERAL INFORMATION:
; APPLICANT: Sigridur Hjorleifsdottir
; APPLICANT: Gudmundur O. Hrengvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aeyarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,875
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Bacteriophage RB69
US-10-270-875-37

```

```

Query Match      60.3%; Score 41; DB 14; Length 898;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      2 PTEILKLVNKRSE 14
      |||:|:~:~:~:~:~
DB      472 PTEITKVFQNRKE 484

```

```

RESULT 12
US-10-270-878-37
; Sequence 37, Application US/10270878
; Publication No. US20030082790A1
; GENERAL INFORMATION:
; APPLICANT: Sigridur Hjorleifsdottir
; APPLICANT: Gudmundur O. Hrengvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aeyarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,878
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Bacteriophage RB69
US-10-270-878-37

```

Query Match 60.3%; Score 41; DB 14; Length 898;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 PTEIACLVNRKSE 14
|||:|:|:|:
Db 472 PTEITKVFNRKKE 484

RESULT 13

US-10-270-786-37
; Sequence 37, Application US/10270786
; Publication No. US20030087392A1
; GENERAL INFORMATION:
; APPLICANT: Sigridur Hjorleifsdottir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Bacteriophage RB69
US-10-270-786-37

Query Match 60.3%; Score 41; DB 14; Length 898;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 PTEIACLVNRKSE 14
|||:|:|:|:
Db 472 PTEITKVFNRKKE 484

RESULT 14

US-10-270-710-37
; Sequence 37, Application US/10270710
; Publication No. US20030092128A1
; GENERAL INFORMATION:
; APPLICANT: Sigridur Hjorleifsdottir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Bacteriophage RB69
US-10-270-710-37

Query Match 60.3%; Score 41; DB 14; Length 898;

Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 PTEIACLVNRKSE 14
|||:|:~|:|:
Db 472 PTEITKVFNRKKE 484

RESULT 15

US-10-270-859-37
; Sequence 37, Application US/10270859
; Publication No. US20030092134A1
; GENERAL INFORMATION:
; APPLICANT: Sigridur Hjorleifsdottir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Bacteriophage RB69
US-10-270-859-37

Query Match 60.3%; Score 41; DB 14; Length 898;
Best Local Similarity 53.8%; Pred. No. 1.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 PTEIACLVNRKSE 14
|||:|:~|:|:
Db 472 PTEITKVFNRKKE 484

Search completed: May 21, 2004, 18:05:12
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 17:59:52 ; Search time 21 Seconds
(without alignments)
64.128 Million cell updates/sec

Title: US-10-045-673a-1

Sequence: 1 TPTELAKLVNKRSE 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR 78:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	86.8	474	1	VYHND
2	44	64.7	314	1	UC4661
3	41	60.3	248	2	AF1851
4	40	58.8	392	2	AF3445
5	40	58.8	593	2	A96783
6	40	58.8	626	2	D86601
7	40	58.8	1124	2	B84742
8	39	57.4	71	2	T44732
9	39	57.4	117	2	S52233
10	39	57.4	622	2	T22716
11	39	57.4	699	2	F95146
12	39	57.4	701	2	D98014
13	39	57.4	710	2	P86778
14	39	57.4	1034	2	A95262
15	38	55.9	255	2	T17866
16	38	55.9	327	2	D72257
17	38	55.9	430	2	C95300
18	38	55.9	555	2	T45351
19	38	55.9	556	2	JC5312
20	38	55.9	723	2	H82035
21	38	55.9	729	1	A39592
22	38	55.9	729	2	P19125
23	38	55.9	729	2	D86072
24	38	55.9	729	2	AH0914
25	37	54.4	251	2	A44506
26	37	54.4	251	2	G90015
27	37	54.4	311	2	B69191
28	37	54.4	339	2	A97669
29	37	54.4	364	2	S26448

30	37	54.4	364	2	S26458	hypothetical prote
31	37	54.4	365	2	AF2893	glycosyltransferas
32	37	54.4	369	2	T50820	hypothetical prote
33	37	54.4	448	2	F84780	hypothetical prote
34	37	54.4	539	2	A84534	hypothetical prote
35	37	54.4	540	2	A86020	hypothetical prote
36	37	54.4	540	2	S47708	hypothetical 61.2K
37	37	54.4	543	2	A98174	hypothetical prote
38	37	54.4	540	2	A46601	saratoxin precu
39	37	54.4	644	2	A16325	epidermal growth f
40	37	54.4	644	2	T47835	hypothetical prote
41	37	54.4	692	2	AC1234	DNA topoisomerase
42	37	54.4	692	2	A11596	DNA topoisomerase
43	37	54.4	819	2	A82549	phage-related prot
44	37	54.4	845	2	G82773	phage-related prot
45	37	54.4	1210	2	A53183	epidermal growth f

ALIGNMENTS

RESULT 1

vitamin D-binding protein precursor [validated] - human
N:Alternate names: DBP; Gc-globulin; group-specific component
C:Date: 28-May-1986 #sequence revision 28-May-1986 #text change 17-Mar-2000
C:Accession: A94076; A46759; A29096; A92765; S39787; A24066; A90427; A03237
R:Yang, F.; Brune, J.L.; Naylor, S.L.; Cupples, R.L.; Naberhaus, K.H.; Bowman, B.H.
Proc. Natl. Acad. Sci. U.S.A. 82, 7994-7998, 1985
A:Title: Human group-specific component (Gc) is a member of the albumin family.
A:Reference number: A94076; MUID:86068030; PMID:2415977
A:Accession: A94076
A:Molecule type: mRNA
A:Residues: 1-474 <YANI>
A:Cross-references: GB:X03179; GB:M1321; NID:G31675; PIDN:CAA26938.1; PID:G31676
A:Experimental source: allele Gc2
R:Witte, W.F.; Glubs, P.E.M.; Zielinski, R.; Yang, F.; Bowman, B.H.; Dugalczyk, A.
Genomics 16, 751-754, 1993
A:Title: Complete structure of the human Gc gene: differences and similarities between
A:Reference number: A46759; MUID:93315171; PMID:8325650
A:Accession: A46759
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-431, 'E', 433-435, 'T', 437-474 <WIR>
A:Cross-references: GB:L10641; NID:G340281; PIDN:AAA61704.1; PID:G639896
A:Experimental source: allele Gc1
R:Yang, F.; Naberhaus, K.H.; Aditman, G.S.; Gardella, J.M.; Brissenden, J.E.; Bowman, B.
Gene 54, 285-290, 1987
A:Title: The vitamin D-binding protein gene contains conserved nucleotide sequences tha
A:Reference number: A29096; MUID:88005794; PMID:2958390
A:Accession: A29096
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-19 <YAN2>
A:Cross-references: GB:M17156; NID:G181489; PIDN:AAA19662.1; PID:G463096
R:Cooke, N.R.; David, E.V.
J. Clin. Invest. 76, 2420-2424, 1985
A:Title: Serum vitamin D-binding protein is a third member of the albumin and alpha fet
A:Reference number: A92765; MUID:86086396; PMID:2416779
A:Accession: A92765
A:Molecule type: mRNA
A:Residues: 1-167, 'E', 169-326, 'R', 328-431, 'E', 433-435, 'T', 437-474 <COO>
A:Cross-references: GB:M12654; NID:G181481; PIDN:AAA52173.1; PID:G181482
R:Brann, A.; Kotler, A.; Morawietz, S.; Cleve, H.
Biochim. Biophys. Acta 1216, 385-394, 1993
A:Title: Sequence and organization of the human vitamin D-binding protein gene.
A:Reference number: S39787; MUID:94092730; PMID:7505619
A:Accession: S39787
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-431, 'E', 433-435, 'T', 437-474 <BRA>

A/Cross-references: GB:567480; NID:9455967; PIDN:AB29423.1; PID:9455970
R/Schoentgen, F.; Metz-Boutigue, M.H.; Jolles, J.; Constans, J.; Jolles, P.
Biochim. Biophys. Acta 871, 189-198, 1986
A/Title: Complete amino acid sequence of human vitamin D-binding protein (group-specific
A/Reference number: A24066; MUID:86216223; PMID:2423133
A/Accession: A24066
A/Molecule type: protein
A/Residues: 17-474 <SCH>
R/Saast, J.; Kurosky, A.; Bennett, A.; Bowman, B.H.
Biochemistry 18, 1611-1617, 1979
A/Title: Molecular basis for the three major forms of human serum vitamin D binding prot
A/Reference number: A90427; MUID:79145448; PMID:218624
A/Accession: A90427
A/Molecule type: protein
A/Residues: 17, 'Q', 19, 'N', 23-36, 'XXX', 40-41, 472-474 <SYA>
C/Comment: DBP is a multifunctional protein found in plasma, ascitic fluid, cerebrospine
nts polymerization of actin by binding its monomers. DBP associates with membrane-bound
C/Comment: Over 80 variants of human DBP have been identified. The three most common all
C/Genetics:
A/Gene: GDB:GC
A/Cross-references: GDB:119263; OMIM:139200
A/Map position: 4q12-4q13
A/Intons: 20/1; 43/2; 87/3; 158/2; 202/3; 234/2; 277/3; 345/2; 388/3; 421/2; 465/3
C/Superfamily: serum albumin; serum albumin repeat homology
C/Keywords: actin binding; duplication; globulin; glycoprotein; liver; plasma; polymorph
F.1-16/Domain: signal sequence #status predicted <SIG>
F.1-474/Product: vitamin D-binding protein (allele Gc2) #status experimental <MAL2>
F.17-431, 'E', 433-435, 'T', 437-474/Product: vitamin D-binding protein (allele Gc1) #status
F.26-199/Domain: serum albumin repeat homology <SA1>
F.217-385/Domain: serum albumin repeat homology <SA2>
F.404-474/Domain: serum albumin repeat homology #status atypical <SA3>
F.29-75, 74-83, 96-112, 111-122, 145-190, 189-198, 220-266, 265-273, 286-300, 299-311, 335-376, 375
F.288/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 86.8%; Score 59; DB 1; Length 474;
Best Local Similarity 85.7%; Pred. No. 0.013;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPTELAKLVNKRSE 14
|||:|||||:
DB 434 TPTELAKLVNKRSD 447

RESULT 2
JC4661
proline dehydrogenase (EC 1.5.99.8) - Photobacterium leiognathi
N/Alternate names: putA protein
C/Species: Photobacterium leiognathi
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C/Accession: JC4661
R/Liu, J.; Yu, K.; Chen, H.; Weng, S.
Biochem. Biophys. Res. Commun. 219, 868-875, 1996
A/Title: Regulatory region with putA gene of proline dehydrogenase that links to the lum
A/Reference number: JC4661; MUID:96216743; PMID:8645272
A/Accession: JC4661
A/Molecule type: DNA
A/Residues: 1-314 <Liu>
A/Cross-references: GB:U93227; NID:91236800; PIDN:AA043866.1; PID:91236801
A/Experimental source: PL741
C/Comment: This enzyme is a multifunctional membrane-associated dehydrogenase.
C/Genetics:
A/Gene: putA
C/Superfamily: proline dehydrogenase
C/Keywords: oxidoreductase

Query Match 64.7%; Score 44; DB 1; Length 314;
Best Local Similarity 64.3%; Pred. No. 4.2;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 TPTELAKLVNKRSE 14
|||:|||||:
DB 154 TPTELAKLVNKRSE 167

RESULT 3
AF1851
hypothetical protein alr0359 [imported] - Nostoc sp. (strain PCC 7120)
C/Species: Nostoc sp. PCC 7120
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C/Accession: AF1851
R/Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriqul
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
A/Reference number: AB1807; MUID:21595285; PMID:11759840
A/Accession: AF1851
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-248 <Kur>
A/Cross-references: GB:BA000019; PIDN:BA072317.1; PID:917129704; GSPDB:GN00179
A/Experimental source: strain PCC 7120
C/Genetics:
A/Gene: alr0359

Query Match 60.3%; Score 41; DB 2; Length 248;
Best Local Similarity 72.7%; Pred. No. 11;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PTELAKLVNKR 12
|||:|||||:
DB 225 PTELAKLVNKR 235

RESULT 4
AF3445
spore-cortex-lytic enzyme prepeptide precursor [imported] - Brucella melitensis (strai
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C/Accession: AF3445
R/DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova
M.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Lete
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen Brucella melite
A/Reference number: AD3252; PMID:11756688
A/Accession: AF3445
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-392 <Kur>
A/Cross-references: GB:AE008917; PIDN:JAL52729.1; PID:917983559; GSPDB:GN00190
A/Experimental source: strain 16M
C/Genetics:
A/Gene: BME11548
A/Map position: 1

Query Match 58.8%; Score 40; DB 2; Length 392;
Best Local Similarity 77.8%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 PTELAKLVN 10
|||:|||||:
DB 190 PTELAKLVN 198

RESULT 5
A96783
unknown protein P22H5.2 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Accession: A96783
R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani;

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11103712
 A:Accession: A96783
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-593 <Sto>
 A:Cross-references: GB:AE005173; NID:g10092269; PIDN:AA612682.1; GSPDB:GM00141
 C:Genetics
 A:Gene: F22H5.2
 A:Map position: 1

Query Match	58.8%	Score 40	DB 2	Length 593
Best Local Similarity	61.5%	Pred. No. 43		
Matches	8	Conservative	2	Mismatches 3
				Indels 0
				Gaps 0
QY	1	TPTELAKLVNRS	13	
Db	85	TPRLAKQVNOOS	97	

```

RESULT 6
D88601
protein Y49E10.11 (imported) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: D88601
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: D88601
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-626 <STO>
A:Cross-references: GB:chr_III; PIDN:CBM1550.1; PID:g3979990; GSPDB:IGN00021; CESP:Y49E10.11
C:Genetics:
A:Gene: Y49E10.11
A:Map position: 3

```

	Query Match	Similarity	Score	DB	Length
Best Local	8%	61.5%	40	2	626
Matches	8	Conservative	2	Mismatches	3
				Indels	0
				Gaps	0

```

RESULT 7
B84742
probable receptor-like protein kinase [imported] - Arabidopsis thaliana
C.Species: Arabidopsis thaliana (mouse-ear cress)
C.Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C.Accession: B84742
R.Lin: X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
W.; Koo, H.; Motilal, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Taitou, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-766, 1999
A.Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A.Reference number: AB4420; MUID:20083487; PMID:10617197
A.Accession: B84742
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1124 <SNO>
A.Cross-references: GB:AE002093; NID:g2924777; PIDN:AAC04906.1; GSPDB:GM00139
C.Genetics:
A.Gene: At2g33170
A.Map position: 2

```

Query Match	58.8%	Score 40	DB 2	Length 1124
Best Local Similarity	88.9%	Pred. No. 83		
Matches	8	Conservative	0	Mismatches 1
				Indels 0
				Gaps 0
QY	2	PTETLAKVYN	10	
DB	486	PTETLCKVYN	494	

RESULT 8
 T44732
 hypothetical protein MLCB1450.16 [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
 C:Accession: T44732
 R:James, K.D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, January 1998
 A:Reference number: Z22831
 A:Accession: T44732
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-71 <JAM>
 A:Cross-references: EMBL:AL035159; PIDD:CAA22700.1
 A:Experimental source: cosmid B1450
 C:Genetics:
 A>Note: MLCB1450.16

Query Match	57.4%	Score 39;	DB 2;	Length 71;
Best Local Similarity	57.1%	Pred. No. 6.9;		
Matches	8;	Conservative	3;	Mismatches
			3;	Indels
			0;	Gaps
0%	1	TPPETAADVNKRS	14	
		: :		
Db	29	TPPMAKTLNRYVE	42	

RESULT 9
 S52233
 hypothetical protein 117 - *Coxiella burnetii*
 C|Species: *Coxiella burnetii*
 C|Date: 05-May-1995 #sequence_revision 01-Sep-1995 #text_change 08-Oct-1999
 C|Accession: S52233; S38221
 R|Willems, H.; Thiele, D.; Valkova, D.
 submitted to the EMBL Data Library, March 1995
 A|Reference number: S52227
 A|Accession: S52233
 A|Molecule type: DNA
 A|Residues: 1-117 <MTL>
 A|Cross-references: EMBL:X85964; NID:g757760; PIDD:CA55946.1; PIR:g757767
 R|Thiele, D.; Willems, H.; Haas, M.; Kraus, H.
 submitted to the EMBL Data Library, October 1993
 A|Reference number: S38215
 A|Accession: S38221
 A|Molecule type: DNA
 A|Residues: 1-117 <MTL>
 A|Cross-references: EMBL:X75356; NID:g407370; PIDD:CA53109.1; PIR:g407377
 C|SpeciesFamily: *Coxiella burnetii* hypothetical protein 117

Query Match	57.4%	Score 39;	DB 2;	Length 117;
Best Local Similarity	57.1%	Pred. No. 12;		
Matches	8;	Conservative	1;	Mismatches 5; Indels 0; Gaps 0;
QY	1	TPTEIAKIVNRSE	14	
				:
				:
DB	99	TPTRLRLVNNV	112	

RESULT 10
T222716
hypothetical protein F55C_8 - *Caenorhabditis elegans*
C/Species: *Caenorhabditis elegans*
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22716
R:Harris, B.
Submitted to the EMBL Data Library, August 1996
A:Reference number: Z19604
A:Accession: T22716
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-622 <N1>
A:Cross-references: EMBL:Z19604, PIDN:CA801573.1, GSPDB:GN00023, CESP:F55C5.8
A:Experimental source: clone F55C5
C:Genetics:
A:Gene: CESP:F55C5.8
A:Map position: 5
A:Introns: 24/3; 44/2; 147/3; 251/3; 561/3

Query Match 57.4%; Score 39; DB 2; Length 622;
Best Local Similarity 58.3%; Pred. No. 68;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 PTELAUVNK 12
Db 499 TDELAUVNK 510

RESULT 11
F95146
DNA topoisomerase I [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 24-Aug-2001
C:Accession: F95146
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A.; Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: F95146
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-699 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK5367.1; PID:G14972745; GSPDB:GN00164; TIGR:SP4
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SPI263
C:Superfamily: DNA topoisomerase I

Query Match 57.4%; Score 39; DB 2; Length 639;
Best Local Similarity 70.0%; Pred. No. 76;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PTELAUVNK 11
Db 505 PTELAUVNK 514

RESULT 12
D98014
DNA topoisomerase (EC 5.99.1.2) [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C:Accession: D98014
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; F
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAnen, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Vaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: D98014
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-701 <KUR>

A:Cross-references: GB:AE007317; PIDN:AAK9944.1; PID:G15458770; GSPDB:GN00174
C:Genetics:
A:Gene: topA
C:Superfamily: DNA topoisomerase I
C:Keywords: isomerase

Query Match 57.4%; Score 39; DB 2; Length 701;
Best Local Similarity 70.0%; Pred. No. 77;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PTELAUVNK 11
Db 511 PTELAUVNK 520

RESULT 13
F86778
DNA topoisomerase (EC 5.99.1.2) [imported] - Lactococcus lactis subsp. lactis (strain)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: F86778
R:Boletín, A.; Winkler, P.; Mauger, S.; Jallón, O.; Malarme, K.; Weisenbach, J.; Ehr
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: F86778
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-710 <STO>
A:Cross-references: GB:AE005176; PID:G12724201; PIDN:AAK05328.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: topA
C:Superfamily: DNA topoisomerase I
C:Keywords: isomerase; phosphotransferase

Query Match 57.4%; Score 39; DB 2; Length 710;
Best Local Similarity 70.0%; Pred. No. 78;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PTELAUVNK 11
Db 518 PTELAUVNK 527

RESULT 14
A95262
probable formate dehydrogenase (EC 1.2.1.2) alpha chain fdog [imported] - Sinorhizobium
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 17-May-2002
C:Accession: A95262
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: A95262
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1034 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK6469.1; PID:G14523056; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSYMA
R:Galibert, F.; Pinar, T.M.; Long, S.R.; Pulver, A.; Abola, P.; Ampe, F.; Barloy-Hubler
Pela, D.; Chalm, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.T.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
heault, P.; Vandebol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: fdog

A;Genome: plasmid
C;Superfamily: formate dehydrogenase
C;Keywords: oxidoreductase

Query Match 57.4%; Score 39; DB 2; Length 1034;
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

OY      1 TPTELAKLVNKR 13
          |||||:|:
Db      699 TPEELAKEINGRA 711

```

RESULT 15
T17866
hypothetical protein A36L - Chlorella virus PBCV-1
C1:Species: Chlorella virus PBCV-1
C1:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C1:Accession: T17866
R1:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A1:Reference number: Z18806
A1:Accession: T17866
A1:Status: preliminary; translated from GB/EMBL/DBJ.
A1:Molecule type: DNA
A1:Residues: 1-255 <GRA>
A1:Cross-references: EMBL:U42580; NID:G4028696; PIDD:AC96734.1
A1:Experimental source: specific host Chlorella strain NC64A
C1:Genetics:
A1:Note: A36L

Query Match	55.9%	Score 38	DB 2	Length 255
Best Local Similarity	61.5%	Pred. No. 40		
Matches	8	Conservative	1	Mismatches 4
				Indels 0
				Gaps 0

```
QY      2 PTEIAKLWNKRSE 14
          |||||:| |
Db      226 PTEIAKKIVGKFEF 238
```

Search completed: May 21, 2004, 18:03:41
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2004, 17:59:12 ; Search time 11 Seconds

(Without alignments)
66.271 Million cell updates/sec

Title: US-10-045-673A-1

Perfect score: 68

Sequence: 1 TPTETLAKVNRSE 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	86.8	474	1	VTDB_HUMAN P02774 homo sapien
2	41	60.3	903	1	DPOI_BPR69 Q30807 bacterioph
3	40	58.8	357	1	RUVB_BALSO Q8Y236 ralsstonia s
4	40	58.8	527	1	TEB_PYPAB Q8Y240 pyrobaculum
5	39	57.4	622	1	SR68_CABEL Q20822 caenorhabdi
6	39	57.4	710	1	TOP1_LACIA Q09880 lactococcus
7	38	55.9	220	1	TRMB_MYCPE Q86wb6 mycoplasma
8	38	55.9	555	1	PRPB_MYCPE Q33064 mycobacteri
9	38	55.9	729	1	FADB_ECO57 Q8X812 e fatty oxi
10	38	55.9	729	1	FADB_ECO1 P21177 e fatty oxi
11	38	55.9	729	1	FADB_SALTY Q8X366 s fatty oxi
12	38	55.9	729	1	FADB_SALTY Q91615 s fatty oxi
13	37	54.4	185	1	RRR_FIGBR Q84295 wigglewort
14	37	54.4	251	1	LACR_STAM Q99873 staphylococ
15	37	54.4	311	1	LACR_STAM P16644 staphylococ
16	37	54.4	311	1	ECRI_METH Q26780 methanobact
17	37	54.4	352	1	MATV_SAXOP Q33134 saxitraga o
18	37	54.4	364	1	CD6V_METTF P23569 methanobact
19	37	54.4	364	1	CD6Z_METTF P25789 methanobact
20	37	54.4	476	1	YHJF_ECOLI P37627 escherichia
21	37	54.4	540	1	YHJF_ECOLI P13208 atracaspis
22	37	54.4	543	1	SRPX_ATREN Q89016 mus musculu
23	37	54.4	606	1	ABD4_MOUSE Q01279 mus musculu
24	37	54.4	1210	1	EGFR_MOUSE Q92122 candida alb
25	37	54.4	1230	1	YGG2_YEAST P53066 saccharomyc
26	36	52.9	181	1	RRR_THETN Q8Y224 thermomaner
27	36	52.9	247	1	NEP4_LYCBS P12670 lycopersico
28	36	52.9	247	1	NEP4_LYCBS P50702 solanum com
29	36	52.9	250	1	OSB1_SOLCO Q9Z7A2 chlamydia p
30	36	52.9	303	1	3MGA_BACSU P37878 bacillus su
31	36	52.9	340	1	OTC_LACPL Q08822 lactobacilli
32	36	52.9	370	1	OTC_ASPNG P11066 aspergillus
33	36	52.9	370	1	OTC_ASPNG P11066 aspergillus

34	36	52.9	453	1	NH12_CABEL Q21701 caenorhabdi
35	36	52.9	544	1	SRFB_SULSO P9560 sulfolobus
36	36	52.9	654	1	REPC3_YEAST P32349 saccharomyc
37	36	52.9	888	1	YOH0_YEAST P40361 saccharomyc
38	36	52.9	978	1	SIR3_YEAST P06701 saccharomyc
39	36	52.9	1048	1	NMD5_YEAST P46970 saccharomyc
40	36	52.9	1440	1	POLG_JAEVN P14403 genome po
41	36	52.9	3432	1	POLG_JAEV1 P27395 genome po
42	36	52.9	3432	1	POLG_JAEV5 P19110 genome po
43	36	52.9	3432	1	POLG_JAEVJ P32886 genome po
44	35	51.5	149	1	RL9_FUSNN Q8Y110 fusobacteri
45	35	51.5	181	1	Y06E_BPT4 P13311 bacterioph

ALIGNMENTS

RESULT 1	VTDB_HUMAN	STANDARD	PRT	474 AA.
AC	P02774; Q16309; Q16310;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Vitamin D-binding protein precursor (DBP) (Group-specific component) (Gc-globulin) (VDB).			
GN	Gc.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86068030; PubMed=2415977;			
RA	Yang F., Brune J.L., Naylor S.L., Cupples R.L., Nabershaus K.H.,			
RA	Bowman B.H.;			
RT	"Human group-specific component (Gc) is a member of the albumin			
RT	family.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:7994-7998(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86068396; PubMed=2416779;			
RA	Cooke N.E., David E.V.;			
RT	"Serum vitamin D-binding protein is a third member of the albumin and			
RT	alpha fetoprotein gene family.";			
RL	J. Clin. Invest. 76:2420-2424(1985).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94092730; PubMed=7505619;			
RA	Braun A., Kofler A., Morawietz S., Cleve H.;			
RT	"Sequence and organization of the human vitamin D-binding protein			
RT	gene.";			
RL	Biochim. Biophys. Acta 1216:385-394(1993).			
RN	[4]			
RP	SEQUENCE OF 17-474.			
RX	MEDLINE=86216223; PubMed=2423133;			
RA	Schoentgen F., Metz-Boutigue M.-H., Jolles J., Constans J.,			
RA	Jolles P.;			
RT	"Complete amino acid sequence of human vitamin D-binding protein			
RT	(group-specific component): evidence of a three-fold internal			
RT	homology as in serum albumin and alpha-fetoprotein.";			
RL	Biochim. Biophys. Acta 871:189-198(1986).			
RN	[5]			
RP	SEQUENCE OF 17-31 AND 431-441.			
RX	MEDLINE=79145448; PubMed=218624;			
RA	Svavali J., Kurosky A., Bennett A., Bowman B.H.;			
RT	"Molecular basis for the three major forms of human serum vitamin D			
RT	binding protein (group-specific component).";			
RL	Biochemistry 18:1611-1617(1979).			
RN	[6]			
RP	SEQUENCE OF 1-19 FROM N.A.			
RX	MEDLINE=88005794; PubMed=2258390;			
RA	Yang F., Nabershaus K.H., Adrian G.S., Gardella J.M., Briesenden J.E.,			

RA Bowman B.H.;
 RT "The vitamin D-binding protein gene contains conserved nucleotide
 RT sequences that respond to heavy metal, adipocyte and mitotic
 RT signals";
 RL Gene 54:285-290(1987).
 RN [7]
 RN VARIANTS GC*2; GC*1F AND GC*1S.
 RX MEDLINE=92316509; PubMed=1352271;
 RA Braun A., Bichmaier R., Cleve H.;
 RT "Molecular analysis of the gene for the human vitamin-D-binding
 RT protein (group-specific component): allelic differences of the common
 RT genetic GC types";
 RL Hum. Genet. 89:401-406(1992).
 RN [8]
 RP SEQUENCE OF 430-446 FROM N.A., AND VARIANTS GC*2A9 CYS-445 AND GC*1A1
 RP HIS-445.
 RX MEDLINE=95242701; PubMed=7725672;
 RA Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;
 RT "Characterization of mutants of the vitamin-D-binding protein/group
 RT specific component: GC abortifacient (1A1) from Australian aborigines and
 RT South African blacks, and 2A9 from south Germany";
 RL Vox Sang. 68:50-54(1995).
 CC -1- FUNCTION: Multifunctional protein found in plasma, ascitic fluid,
 CC cerebrospinal fluid, and urine and on the surface of many cell
 CC types. In plasma, it carries the vitamin D steroids and prevents
 CC polymerization of actin by binding its monomers. DBP associates
 CC with membrane-bound immunoglobulin on the surface of b-lymphocytes
 CC and with IgG-fc receptor on the membranes of T-lymphocytes.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- POLYMORPHISM: Over 80 variants of human DBP have been identified.
 CC The three most common alleles are called GC*1F, GC*1S, and GC*2.
 CC -1- SIMILARITY: Belongs to the ALB/AFPI/VDB family.
 CC -1- SIMILARITY: Contains 3 albumin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; L10641; AAA61704.1; -;
 DR EMBL; X03178; CAA26938.1; -;
 DR EMBL; M12654; AAA52173.1; -;
 DR EMBL; S67480; AAB29423.1; JOINED.
 DR EMBL; S67474; AAB29423.1; JOINED.
 DR EMBL; S67476; AAB29423.1; JOINED.
 DR EMBL; S67478; AAB29423.1; JOINED.
 DR EMBL; S67479; AAB29423.1; JOINED.
 DR EMBL; S67526; AAB29423.1; JOINED.
 DR EMBL; M17156; AAA19662.2; -;
 DR EMBL; S77129; AAD14249.1; ALT_SEQ.
 DR EMBL; S77130; AAD14250.1; ALT_SEQ.
 DR PIR; A94076; VYRHU.
 DR PDB; 1J78; 06-FEB-02.
 DR PDB; 1J7E; 06-FEB-02.
 DR PDB; 1KW2; 19-JUN-02.
 DR PDB; 1KXP; 19-JUN-02.
 DR PDB; 1LOT; 31-JUL-02.
 DR PDB; 1MA9; 11-FEB-03.
 DR GLYCOsuiteDB; P02774; -;
 DR SWISS-2DPAGE; P02774; HUMAN.
 DR Stena-2DPAGE; P02774; -;
 DR Genew; HGNC:4187; GC.
 DR MIM; 139200; -;
 DR GO; GO:0005386; F:carrier activity; TAS.
 DR GO; GO:0005499; F:vitamin D binding; TAS.
 DR GO; GO:0015875; P:vitamin/cofactor transport; TAS.
 DR InterPro; IPR000264; Serum_albumin.
 DR Pfam; PF00273; transport_prot; 2.
 DR PRINTS; PR00802; SERUMALBUMIN.

DR ProDom; PD002486; Serum_albumin; 1.
 DR SMART; SM00103; ALBUMIN; 2.
 DR PROSITE; PS00212; ALBUMIN; 1.
 KM Glycoprotein; Vitamin D; Transport; Actin-binding; Repeat;
 KW Polymorphism; Signal; 3D-structure.
 FT SIGNAL 1 16
 FT CHAIN 17 474
 FT DOMAIN 20 202
 FT DOMAIN 208 388
 FT DOMAIN 395 474
 FT DISULFID 29 75
 FT DISULFID 74 83
 FT DISULFID 96 112
 FT DISULFID 111 122
 FT DISULFID 145 190
 FT DISULFID 189 198
 FT DISULFID 220 266
 FT DISULFID 265 273
 FT DISULFID 286 300
 FT DISULFID 299 311
 FT DISULFID 335 376
 FT DISULFID 375 384
 FT DISULFID 407 453
 FT DISULFID 452 462
 FT CARBOHYD 288
 FT VARIANT 432
 FT VARIANT 436
 FT VARIANT 445
 FT VARIANT 445
 FT VARIANT 445
 FT CONFLICT 168 168
 FT CONFLICT 327 327
 FT SEQUENCE 474 AA; 52963 MW; 6AD8F16B551F1E4 CRC64;
 SQ
 Query Match 86.8%; Score 59; DB 1; Length 474;
 Best Local Similarity 85.7%; Pred. No. 0.0053;
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TPPELALVNRKSE 14
 Db 434 TPPELALVNRKSD 447
 AC Q38087;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE DNA polymerase (EC 2.7.7.7) (Gp43).
 GN 43.
 OS Bacteriophage RB69.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 CX NCBI_TaxID=12353;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9604739; PubMed=7592876;
 RA Wang C.C., Yeh L.-S., Karam J.D.;
 RT "Molecular organization of T4 DNA polymerase. Evidence from
 RT phylogenetics";
 RL J. Biol. Chem. 270:26558-26564(1995).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=97358535; PubMed=9215631;
 RA Wang J., Sattar A.K., Karam J.D., Konigsberg W.H.,
 RA Steitz T.A.;
 RT "Crystal structure of a pol alpha family replication DNA polymerase

RT from bacteriophage RB69."
RL Cell 89:1087-1093(1997).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF COMPLEX WITH GP45.
RX MEDLINE=20004392; PubMed=10535734;
RA Shamooy, Steitz T.A.
RT "Building a replisome from interacting pieces: sliding clamp
RT complexed to a peptide from DNA polymerase and a polymerase editing
RT complex."
RL Cell 99:155-166(1999).
CC -1- FUNCTION: This polymerase possesses two enzymatic activities: DNA
CC synthesis (polymerase) and an exonucleolytic activity that
CC degrades single stranded DNA in the 3' to 5' direction.
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA} (N).
CC -1- SIMILARITY: Belongs to the DNA polymerase type-B family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U34036; AAA93077.1; -.
DR PDB; 1WAF; 14-JAN-98.
DR PDB; 1WAF; 14-JAN-98.
DR PDB; 1CLQ; 28-OCT-99.
DR PDB; 1B1F; 29-DEC-99.
DR PDB; 1B8H; 05-NOV-99.
DR PDB; 1IG9; 11-JUN-01.
DR PDB; 1IH7; 13-JUN-01.
DR PDB; 1WAG; 14-JAN-98.
DR PDB; 1WAF; 14-JAN-98.
DR InterPro; IPR006172; DNA_pol_B.
DR InterPro; IPR006134; DNA_pol_B_dom.
DR InterPro; IPR006133; DNA_pol_B_exo.
DR Pfam; PF00136; DNA_pol_B; 1.
DR Pfam; PF03104; DNA_pol_B_exo; 1.
DR PRINTS; PRO0106; DNAPOLB.
DR SMART; SM00486; POLB; 1.
DR PROSITE; PS00116; DNA_POLYMERASE_B; 1.
DR Transferrase; DNA-directed DNA polymerase; DNA replication; Hydrolyase;
KW Exonuclease; DNA-binding; 3D-structure.
KV STRAND 4 11
FT TURN 12 13
FT STRAND 14 20
FT TURN 22 23
FT STRAND 26 31
FT STRAND 36 40
FT TURN 43 44
FT STRAND 50 51
FT TURN 52 53
FT STRAND 56 61
FT STRAND 65 78
FT HELIX 79 79
FT TURN 83 83
FT STRAND 88 96
FT HELIX 105 107
FT STRAND 110 116
FT TURN 126 128
FT STRAND 135 140
FT STRAND 141 144
FT TURN 145 150
FT STRAND 153 154
FT STRAND 155 156
FT TURN 157 158
FT STRAND 164 167
FT TURN 168 168
FT HELIX 171 173
FT TURN 174 174
FT HELIX 180 183

FT TURN 184 185
FT STRAND 186 190
FT HELIX 194 207
FT TURN 208 208
FT STRAND 212 214
FT TURN 218 221
FT HELIX 222 234
FT TURN 236 239
FT HELIX 240 242
FT TURN 244 245
FT STRAND 248 255
FT TURN 256 257
FT STRAND 258 265
FT TURN 266 267
FT STRAND 269 270
FT HELIX 273 280
FT HELIX 290 298
FT HELIX 309 311
FT HELIX 312 315
FT HELIX 317 338
FT HELIX 340 351
FT TURN 352 352
FT HELIX 355 359
FT HELIX 361 374
FT TURN 375 377
FT STRAND 378 379
FT STRAND 402 403
FT STRAND 407 412
FT TURN 413 414
FT HELIX 415 423
FT TURN 424 424
FT HELIX 427 429
FT STRAND 430 433
FT HELIX 439 443
FT TURN 444 445
FT STRAND 454 456
FT TURN 458 459
FT STRAND 461 464
FT HELIX 470 503
FT TURN 504 504
FT TURN 518 519
FT HELIX 524 530
FT TURN 531 532
FT HELIX 535 570
FT TURN 571 571
FT TURN 573 574
FT TURN 576 577
FT HELIX 580 608
FT TURN 609 610
FT STRAND 617 620
FT STRAND 624 627
FT TURN 629 629
FT HELIX 630 634
FT TURN 635 639
FT HELIX 644 657
FT TURN 658 658
FT HELIX 659 673
FT TURN 674 675
FT STRAND 683 689
FT TURN 693 694
FT STRAND 700 704
FT TURN 705 706
FT STRAND 707 715
FT TURN 716 717
FT STRAND 718 728
FT TURN 732 733
FT TURN 735 736
FT HELIX 739 754
FT TURN 755 755
FT HELIX 757 772
FT TURN 773 773
FT HELIX 776 779
FT STRAND 781 784

```

FT HELIX 788 791
FT STRAND 792 793
FT TURN 794 795
FT STRAND 796 797
FT TURN 799 800
FT HELIX 803 814
FT TURN 826 827
FT STRAND 829 835
FT TURN 837 838
FT TURN 840 841
FT STRAND 845 849
FT TURN 856 858
FT HELIX 859 865
FT STRAND 866 866
FT HELIX 868 875
FT TURN 876 876
FT HELIX 877 887
FT TURN 888 888
FT TURN 897 900
SQ SEQUENCE 903 AA; 104613 MW; A3983FCL6D4C0509 CRC64;

Query Match 60.3%; Score 41; DB 1; Length 903;
Best Local Similarity 53.8%; Pred. NO. 18;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PTELAKLVNKRSE 14
Db 472 PTEITVFNORKE 484

RESULT 3
RUVB_RALSO STANDARD; PRT; 357 AA.
AC 08Y236;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Holliday junction DNA helicase RUVB.
GN RUVB OR RSC0500 OR RSO5022.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OC NCBI_TaxID=305;
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mangelot S.,
RA Ariat M., Billault A., Broctier P., Camus J.C., Catolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebaud P., Whalen M., Winkler P., Levy M.,
RA Weisenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RT Nature 415:497-502(2002).
CC -1- FUNCTION: The ruva-ruvb complex in the presence of ATP renatures
CC cruciform structure in supercoiled DNA with palindromic sequence,
CC indicating that it may promote strand exchange reactions in
CC homologous recombination. RUVB is an helicase that mediates the
CC Holliday junction migration by localized denaturation and
CC reannealing (By similarity).
CC -1- SUBUNIT: Forms a complex with ruva (By similarity).
CC -1- SIMILARITY: Belongs to the ruvb family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL64059; CAD14028.1; -

```

```

DR HAMAP; MF_00016; -; 1.
DR InterPro; IPR003593; AAA_Arase.
DR InterPro; IPR003959; AAA_Arase_cent.
DR InterPro; IPR004605; RuvB.
DR InterPro; IPR008823; RuvB_C.
DR InterPro; IPR008824; RuvB_N.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF05491; RuvB_C; 1.
DR Pfam; PF05496; RuvB_N; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00635; ruvb; 1.
DR DNA repair; SOS response; ATP-binding; DNA recombination; Helicase;
KW Complete proteome.
FT NE_BIND 73
FT BLIND 80
SQ SEQUENCE 357 AA; 39058 MW; 47795DB5F7AD7DC6 CRC64;

Query Match 58.8%; Score 40; DB 1; Length 357;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPTELAKLVNKR 11
Db 196 TPTELARIYTR 206

RESULT 4
TP6B_PYRAB STANDARD; PRT; 527 AA.
AC 08ZVW0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Type II DNA topoisomerase VI subunit B (EC 5.99.1.3) (TopoVI-B).
GN TOP6B OR PAE2217.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OC NCBI_TaxID=13773;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RC MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Lader H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum";
RT Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -1- FUNCTION: Relaxes both positive and negative superturns and
CC exhibits a strong decatenase activity. The B subunit binds ATP (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: Heterotetramer of two subunits A and two subunits B (By
CC similarity).
CC -1- SIMILARITY: Belongs to the TOP6B family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009863; AAP64036.1; -
DR HAMAP; MF_00322; -; 1.
DR InterPro; IPR003594; ATPbind_Arase.
DR InterPro; IPR005734; DNA_top6B.
DR Pfam; PF02518; HATPase_c; 1.
DR SMART; SM00387; HATPase_c; 1.
DR TIGRFAMs; TIGR01052; top6b; 1.
KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Complete proteome.
SQ SEQUENCE 527 AA; 59870 MW; 7BADB81F8C9D97 CRC64;

```

Query Match 58.8%; Score 40; DB 1; Length 527;
 Best Local Similarity 57.1%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 TPTELATLVNKRSE 14
 |||:|||||:
 Db 284 TPTELATLVNKRKQ 297

RESULT 5
 SR68 CAEEL STANDARD; PRT; 622 AA.
 ID SR68 CAEEL STANDARD; PRT; 622 AA.
 AC Q20822;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Probable signal recognition particle 68 kDa protein (SRP68).
 GN F55C5.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Harris B.;
 RU Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Signal-recognition-particle assembly has a crucial role
 in targeting secretory proteins to the rough endoplasmic reticulum
 membrane. SRP68 binds the 7S RNA. SRP72 binds to this complex
 subsequently. This ribonucleoprotein complex might interact
 directly with the docking protein in the ER membrane and possibly
 participate in the elongation arrest function (By similarity).
 CC -1- SUBUNIT: Signal recognition particle consists of a 7S RNA molecule
 of 300 nucleotides and six protein subunits: SRP72, SRP68, SRP54,
 SRP19, SRP14 and SRP9 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the SRP68 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; Z76198; CAB01573.1; -
 DR PIR; T22716; T22716.
 DR Wormpep; F55C5.8; CE20875.
 DR InterPro; IPR008941; TPR-1like.
 KW Hypothetical protein; Signal recognition particle; Ribonucleoprotein;
 KM RNA-binding.
 SQ SEQUENCE 622 AA; 70574 MW; A7B8808E46169636 CRC64;

Query Match 57.4%; Score 39; DB 1; Length 622;
 Best Local Similarity 56.3%; Pred. No. 29;
 Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 TPTELATLVNKR 12
 |||:|||||:
 Db 499 TDSLATKIDKR 510

RESULT 6
 TOP1 LACIA STANDARD; PRT; 710 AA.
 ID TOP1 LACIA STANDARD; PRT; 710 AA.
 AC Q9CG80;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA topoisomerase I (EC 5.99.1.2) (Omega-protein) (Relaxing enzyme)

DN (Unwisting enzyme) (Swivelase).
 GN TOPA OR IL1230.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235166; PubMed=11337471;
 RA Bolotin A., Wincker P., Manger S., Jallou O., Malarne K.,
 RA Weisenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 lactis ssp. lactis IL1403."
 RL Genome Res. 11:731-753(2001).
 CC -1- FUNCTION: The reaction catalyzed by topoisomerases leads to the
 conversion of one topological isomer of DNA to another.
 CC -1- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
 DNA, followed by passage and rejoining.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- MISCELLANEOUS: When a topoisomerase transiently breaks a DNA
 backbone bond, it simultaneously forms a protein-DNA link, in
 which a tyrosyl oxygen in the enzyme is joined to a DNA phosphorus
 at one end of the enzyme-severed DNA strand.
 CC -1- SIMILARITY: Belongs to the prokaryotic type I/III topoisomerase
 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; AE006355; AAC05328.1; -
 DR PIR; F86778; F86778.
 DR HSP; P06612; IECL.
 DR InterPro; IPR005733; DNA_topi_bact.
 DR InterPro; IPR000380; DNA_topiase.
 DR InterPro; IPR003601; DNATopi_ATP_bind.
 DR InterPro; IPR003602; DNATopi_DNA_bind.
 DR InterPro; IPR006171; Toprim_dom.
 DR InterPro; IPR006154; Toprim_sub.
 DR Pfam; PF01131; Topoisom_bac; 1.
 DR Pfam; PF01751; Toprim; 1.
 DR Pfam; PF01386; zf-C4_Topoism; 3.
 DR PRINTS; PR00417; PRTISMASE1.
 DR SMART; SM00437; TOPIAC; 1.
 DR SMART; SM00436; TOPIBC; 1.
 DR SMART; SM00493; TOPRIM; 1.
 DR TIGRFAMs; TIGR01051; TOPA_bact; 1.
 DR PROSITE; PS00396; TOPOISOMERASE_I_PROK; 1.
 KW Isomerase; Topoisomerase; DNA-binding; Zinc-finger; Metal-binding;
 KM Repeat; Complete proteome.
 FT ZN_FING 595 621 C4-TYPE 1.
 FT ZN_FING 635 663 C4-TYPE 2.
 FT ZN_FING 676 702 C4-TYPE 3 (ATYPICAL).
 FT ACT_SITE 321 321 DNA_CLEAVAGE (BY SIMILARITY).
 SQ SEQUENCE 710 AA; 80741 MW; 1268B706C38A59EB CRC64;

Query Match 57.4%; Score 39; DB 1; Length 710;
 Best Local Similarity 70.0%; Pred. No. 34;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 PTELATLVNKR 11
 |||:|||||:
 Db 518 PTELATLVNKR 527

RESULT 7
 TRMB MYCPE STANDARD; PRT; 220 AA.
 ID TRMB MYCPE STANDARD; PRT; 220 AA.
 AC Q8EWB6;

DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE tRNA (guanine-N(7)-)-methyltransferase (EC 2.1.1.33) (tRNA (m7G46) - methyltransferase).
 GN MYP2880.
 OS Mycoplasma penetrans.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=28227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HF-2;
 RX MEDLINE=22354719; PubMed=1246555;
 RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K., Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
 RT "The complete genomic sequence of Mycoplasma penetrans, an intracellular bacterial pathogen in humans.";
 RL Nucleic Acids Res. 30:5293-5300(2002).
 CC -1- FUNCTION: Catalyzes the formation of N(7)-methylguanine at position 46 (m7G46) in tRNA (By similarity).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-homocysteine + tRNA containing N(7)-methylguanine.
 CC -1- SIMILARITY: Belongs to the methyltransferase superfamily. Trmb family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF004171; BAC4080.1; -.
 DR HAMAP; MF 01057; -; 1.
 DR InterPro; IPR004395; Cons_hypoth91.
 DR InterPro; IPR003358; Methyltransf_4.
 DR Pfam; PF02390; Methyltransf_4; 1.
 DR TrRPMs; TIGR00091; TIGR0091; 1.
 DR Transferrase; Methyltransferase; tRNA processing; Complete proteome.
 SQ SEQUENCE 220 AA; 25888 MW; FFA0D7E5F19DA310 CRC64;
 QY Query Match 55.9%; Score 38; DB 1; Length 220;
 Best Local Similarity 80.0%; Pred. No. 16;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Db 2 PTEIAKLVNK 11
 74 PTEIAKLVNK 83
 RESULT 8
 ID FPRB_MYCLE STANDARD; PRT; 555 AA.
 AC 03064;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable ferredoxin/ferredoxin-NADP reductase (EC 1.1.18.1.2) (FNR).
 GN FPRB OR ML2134 OR MCB57.39.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R., Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D., Mungall K., Basham D., Brown D., Chillingworth T., Connor R., Davies R.M., Devlin K., Duthoy S., Felwell T., Fraser A., Hamlin N., Holtroyd S., Hornby T., Jaseis K., Lacroix C., Maclean J., Moule S.,

RA Murphy L., Oliver K., Quail M.A., Rajendram M.A., Rutherford K.M., Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R., RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R., RA Barelli B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -1- CATALYTIC ACTIVITY: Reduced ferredoxin + NADP(+) = oxidized ferredoxin + NADPH.
 CC -1- COFACTOR: FAD; probably binds one or two 4Fe-4S clusters.
 CC -1- SIMILARITY: In the N-terminal, belongs to the bacterial-type ferredoxin family.
 CC -1- SIMILARITY: IN THE C-TERMINAL, TO OTHER FERREDOXIN NADP REDUCTASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z99494; CAB1679.1; -.
 DR EMBL; AL583924; CAC31089.1; -.
 DR PIR; T45351; T45351.
 DR HSSP; P55907; IXER.
 DR Lepidoma; ML2134; -.
 DR InterPro; IPR001450; 4Fe4S_ferredoxin.
 DR InterPro; IPR000813; 7Fe_ferredoxin.
 DR InterPro; IPR000759; Adrxr reductase.
 DR InterPro; IPR001327; FAD_pyr_redox.
 DR Pfam; PF00037; fer4; 2.
 DR PRINTS; PR00354; 7FESFERDOXIN.
 DR PRINTS; PR00419; ADXRPASE.
 DR PRINTS; PR00368; FADPFR.
 DR PROSITE; PS00198; 4Fe4S_FERREDOXIN; 1.
 KM Oxidoreductase; Flavoprotein; NADP; FAD; Electron transport;
 KW Iron-sulfur; 4Fe-4S; Complete proteome.
 FT DOMAIN 1 83 FERREDOXIN.
 FT 115 555 FERREDOXIN-NADP REDUCTASE.
 FT METAL 9 9 IRON-SULFUR 1 (BY SIMILARITY).
 FT METAL 15 15 IRON-SULFUR 1 (BY SIMILARITY).
 FT METAL 19 19 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 45 46 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 49 49 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 52 52 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
 FT METAL 56 56 IRON-SULFUR 1 (BY SIMILARITY).
 SQ SEQUENCE 555 AA; 59712 MW; 43C729286A2DFD9 CRC64;
 QY Query Match 55.9%; Score 38; DB 1; Length 555;
 Best Local Similarity 72.7%; Pred. No. 40;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Db 2 PTEIAKLVNR 12
 492 PTEIAKLVNR 502
 RESULT 9
 ID FADB_ECO57 STANDARD; PRT; 729 AA.
 AC 08X812;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Fatty acid oxidation complex alpha subunit [includes: Enoyl-CoA hydratase (EC 4.2.1.17); Delta(3)-cis-delta(2)-trans-enoyl-CoA isomerase (EC 5.3.3.8); 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35); 3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3)].
 GN FADB OR OLD8 OR Z5367 OR ECS4774.
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.

NCBI_TaxID=83334;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RA MEDLINE=21074935; PubMed=11206551;
 RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Poisel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobleck E.J., Davis N.W., Lim A., Dimalanta E.T., Potanowski K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7";
 RL Nature 409:529-533 (2001).
 [2]
 SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RA MEDLINE=21156231; PubMed=11258796;
 RX Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaoka C., Ogasawara N., Yasunaga T.,
 RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*
 O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22 (2001).
 [3]
 CC -1- FUNCTION: FadB and fadA are the alpha and beta subunits of the
 CC multifunctional enzyme complex of the fatty acid degradation
 CC cycle (by similarity).
 CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxyacyl-CoA + NAD(+) = 3-oxoacyl-CoA
 CC + NADH.
 CC -1- CATALYTIC ACTIVITY: (3S)-3-hydroxybutanoyl-CoA = trans-2(or 3)-enoyl-
 CC CoA + H(2)O.
 CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxybutanoyl-CoA = (R)-3-
 CC hydroxybutanoyl-CoA.
 CC -1- CATALYTIC ACTIVITY: 3-cis-dodecenoyl-CoA = 2-trans-dodecenoyl-CoA.
 CC -1- PATHWAY: Fatty acid beta-oxidation cycle; third step.
 CC -1- SUBUNIT: Tetramer of two alpha chains and two beta chains (by
 CC similarity).
 CC -1- SIMILARITY: In the N-terminal section; belongs to the 3-
 CC hydroxyacyl-CoA dehydrogenase family.
 CC -1- SIMILARITY: In the C-terminal section; belongs to the enoyl-CoA
 CC hydratase/isomerase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL, AB005615; AAG59040.1; -;
 DR EMBL, AP002567; BAB38197.1; -;
 DR PIR, F91225; F91225.
 DR InterPro: IPR006180; 3HCDH.
 DR InterPro: IPR006108; 3HCDH_C.
 DR InterPro: IPR006176; 3HCDH_N.
 DR InterPro: IPR008927; 6OGDH_C like.
 DR InterPro: IPR001753; EnCoA_hydratase.
 DR Pfam: PF00725; 3HCDH; 2.
 DR Pfam: PF02737; 3HCDH_N; 1.
 DR Pfam: PF00378; ECH; 1.
 DR PROSITE: PS00067; 3HCDH; 1.
 DR PROSITE: PS00166; ENOYL-CoA HYDRATASE; 1.
 KM Fatty acid metabolism; Multifunctional enzyme; Oxidoreductase; NAD;
 KM Lyase; Isomerase; Complete proteome.
 SQ SEQUENCE 729 AA; 79559 MW; C13EBBFFPA749183 CRC64;

Query Match 55.9%; Score 38; DB 1; Length 729;
 Best Local Similarity 66.7%; Pred. No. 52;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

3 TELATLVNKRSE 14
 ||| ||| |||

Db 354 TEAAKLANKOLE 365
 RESULT 10
 FADB_ECOLI STANDARD; PRT; 729 AA.
 AC P21177;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Fatty oxidation complex alpha subunit [includes: Enoyl-CoA hydratase
 DE (EC 4.2.1.17); Delta(3)-cis-delta(2)-trans-enoyl-CoA isomerase
 DE (EC 5.3.3.8); 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35); 3-
 DE hydroxybutyryl-CoA epimerase (EC 5.1.2.3)].
 GN FADB OR OLD OR B3846.
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Escherichia*.
 CC NCBI_TaxID=562;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=K12 / MG1655;
 CC MEDLINE=92358234; PubMed=1379743;
 CC Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
 CC "Analysis of the *Escherichia coli* genome: DNA sequence of the region
 CC from 84.5 to 86.5 minutes.";
 CC Science 257:771-778 (1992).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN=K12 / W3110;
 CC MEDLINE=90370500; PubMed=2204034;
 CC Nakahigashi K., Inokuchi H.;
 CC "Nucleotide sequence of the fadA and fadB genes from *Escherichia*
 CC coli.";
 CC Nucleic Acids Res. 18:4937-4937 (1990).
 CC [3]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=91035260; PubMed=1699931;
 CC DiRusso C.C.;
 CC "Primary sequence of the *Escherichia coli* fadB operon, encoding the
 CC fatty acid-oxidizing multienzyme complex, indicates a high degree of
 CC homology to eucaryotic enzymes.";
 CC J. Bacteriol. 172:6459-6468 (1990).
 CC [4]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=91291827; PubMed=1712230;
 CC Yang X.Y.H., Schulz H., Elzinga M., Yang S.Y.;
 CC "Nucleotide sequence of the promoter and fadB gene of the fadB
 CC operon and primary structure of the multifunctional fatty acid
 CC oxidation protein from *Escherichia coli*.";
 CC Biochemistry 30:6788-6795 (1991).
 CC -1- FUNCTION: FadB and fadA are the alpha and beta subunits of the
 CC multifunctional enzyme complex of the fatty acid degradation
 CC cycle.
 CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxyacyl-CoA + NAD(+) = 3-oxoacyl-CoA
 CC + NADH.
 CC -1- CATALYTIC ACTIVITY: (3S)-3-hydroxybutanoyl-CoA = trans-2(or 3)-enoyl-
 CC CoA + H(2)O.
 CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxybutanoyl-CoA = (R)-3-
 CC hydroxybutanoyl-CoA.
 CC -1- CATALYTIC ACTIVITY: 3-cis-dodecenoyl-CoA = 2-trans-dodecenoyl-CoA.
 CC -1- PATHWAY: Fatty acid beta-oxidation cycle; second step.
 CC -1- SUBUNIT: Tetramer of two alpha chains and two beta chains.
 CC -1- SIMILARITY: In the N-terminal section; belongs to the 3-
 CC hydroxyacyl-CoA dehydrogenase family.
 CC -1- SIMILARITY: In the C-terminal section; belongs to the enoyl-CoA
 CC hydratase/isomerase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M87049; AAA67643.1; -

DR EMBL; AE000460; AAC76849.1; -

DR EMBL; X52837; CAB40809.1; -

DR EMBL; M59368; AAA23750.1; -

DR EMBL; M74164; AAA62777.1; -

DR PIR; A39592; A39592.

DR HSSP; P00348; 3HCDH.

DR ECODBASE; G073.4; 6TH EDITION.

DR ECODBASE; G073.5; 6TH EDITION.

DR Ecogene; EGI0279; fadh.

DR InterPro; IPR006180; 3HCDH.

DR InterPro; IPR006108; 3HCDH.C.

DR InterPro; IPR006176; 3HCDH.N.

DR InterPro; IPR008927; 6DGDH.C like.

DR InterPro; IPR001753; EnCoA_hydrtase.

DR Pfam; PF00725; 3HCDH.2.

DR Pfam; PF02737; 3HCDH.N; 1.

DR Pfam; PF00378; ECH; 1.

DR PROSITE; PS00067; 3HCDH; 1.

DR PROSITE; PS00166; ENOYL-CoA HYDRATASE; 1.

KW Fatty acid metabolism; Multifunctional enzyme; Oxidoreductase; NAD; Lyase; Isomerase; Complete proteome.

KM CONFLICT 518 518 A -> R (IN REF. 3).

FT CONFLICT 664 664 F -> L (IN REF. 2).

FT CONFLICT 666 666 P -> A (IN REF. 2).

SO SEQUENCE 729 AA; 79593 MW; 6F1055E402F6B129 CRC64;

Query Match 55.9%; Score 38; DB 1; Length 729;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 TELAKUNTRSE 14
DB 354 TEAAKUNTRSE 365

RESULT 11
FADB_SALTY STANDARD; PRT; 729 AA.

AC 0823C6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fatty oxidation complex alpha subunit [Includes: Enoyl-CoA hydratase (EC 4.2.1.17); Delta(3)-cis-delta(2)-trans-enoyl-CoA isomerase (EC 5.3.3.8); 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35); 3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3)].
GN FADB OR STY3577 OR T3315.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
NCBI_Taxid=601;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J., Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M., Baker S., Basham D., Brooks K., Chillingworth T., Connor P., Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J., Feltham T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagsels K., Krogh A., Larsen T.S., Leach S., Moul S., O'Goora P., Parry C., Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;

RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Iyou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J., Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and CT18.";
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- FUNCTION: Fadh and fadh are the alpha and beta subunits of the multifunctional enzyme complex of the fatty acid degradation cycle (By similarity).
CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxyacyl-CoA + NAD(+) = 3-oxoacyl-CoA + NADH.
CC -1- CATALYTIC ACTIVITY: (3S)-3-hydroxyacyl-CoA = trans-2(or 3)-enoyl-CoA + H(2)O.
CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxybutanoyl-CoA = (R)-3-hydroxybutanoyl-CoA.
CC -1- CATALYTIC ACTIVITY: 3-cis-dodecenoyl-CoA = 2-trans-dodecenoyl-CoA.
CC -1- PATHWAY: Fatty acid beta-oxidation cycle; third step.
CC -1- SUBUNIT: Tetramer of two alpha chains and two beta chains (By similarity).
CC -1- SIMILARITY: In the N-terminal section; belongs to the 3-hydroxyacyl-CoA dehydrogenase family.
CC -1- SIMILARITY: In the C-terminal section; belongs to the enoyl-CoA hydratase/isomerase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AL627278; CAD07910.1; -

DR EMBL; AE016845; AAO70843.1; -

DR InterPro; IPR006180; 3HCDH.

DR InterPro; IPR006108; 3HCDH.C.

DR InterPro; IPR006176; 3HCDH.N.

DR InterPro; IPR008927; 6DGDH.C like.

DR InterPro; IPR001753; EnCoA_hydrtase.

DR Pfam; PF00725; 3HCDH.2.

DR Pfam; PF02737; 3HCDH.N; 1.

DR Pfam; PF00378; ECH; 1.

DR PROSITE; PS00067; 3HCDH; 1.

DR PROSITE; PS00166; ENOYL-CoA HYDRATASE; 1.

KW Fatty acid metabolism; Multifunctional enzyme; Oxidoreductase; NAD; Lyase; Isomerase; Complete proteome.

KM SEQUENCE 729 AA; 79642 MW; C34C6B8CF93713C CRC64;

Query Match 55.9%; Score 38; DB 1; Length 729;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 TELAKUNTRSE 14
DB 354 TEAAKUNTRSE 365

RESULT 12
FADB_SALTY STANDARD; PRT; 729 AA.

AC 0916L5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fatty oxidation complex alpha subunit [Includes: Enoyl-CoA hydratase (EC 4.2.1.17); Delta(3)-cis-delta(2)-trans-enoyl-CoA isomerase (EC 5.3.3.8); 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35); 3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3)].
GN FADB OR STY3583 OR STMD1.6.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
NCBI_Taxid=602;

```

RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwolik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
CC -1- FUNCTION: PadB and fadA are the alpha and beta subunits of the
CC multifunctional enzyme complex of the fatty acid degradation
CC cycle (By similarity).
CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxyacyl-CoA + NAD(+) = 3-oxoacyl-CoA
CC + NADH.
CC -1- CATALYTIC ACTIVITY: (3S)-3-hydroxyacyl-CoA = trans-2(or 3)-enoyl-
CC CoA + H(2O).
CC -1- CATALYTIC ACTIVITY: (S)-3-hydroxybutanoyl-CoA = (R)-3-
CC hydroxybutanoyl-CoA.
CC -1- CATALYTIC ACTIVITY: 3-cis-dodecenoyl-CoA = 2-trans-dodecenoyl-CoA.
CC -1- PATHWAY: Fatty acid beta-oxidation cycle) third step.
CC -1- SUBUNIT: Tetramer of two alpha chains and two beta chains (By
CC similarity).
CC -1- SIMILARITY: In the N-terminal section; belongs to the 3-
CC hydroxyacyl-CoA dehydrogenase family.
CC -1- SIMILARITY: In the C-terminal section; belongs to the enoyl-CoA
CC hydratase/isomerase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF233324; AAF33409.1; -.
DR EMBL; AE008866; AAL22827.1; -.
DR StyGene; SG77777; fadB.
DR HSP; P00348; 3HSD.
DR InterPro; IPR006180; 3HSDH.
DR InterPro; IPR006108; 3HSDH_C.
DR InterPro; IPR006176; 3HSDH_N.
DR InterPro; IPR008927; 6DGDH_C-like.
DR InterPro; IPR001753; EnCoA_hydrtse.
DR Pfam; PF00725; 3HSDH_2.
DR Pfam; PF02737; 3HSDH_N; 1.
DR Pfam; PF00378; ECH; 1.
DR PROSITE; PS00067; 3HSDH_1.
DR PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
KW Fatty acid metabolism; Multifunctional enzyme; Oxidoreductase; NAD;
KW lyase; Isomerase; Complete proteome.
SQ SEQUENCE 729 AA; 79594 MW; 90A6EAB57871582D CRC64;

Query Match 55.9%; Score 38; DB 1; Length 729;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

GN FRR OR WIGBR3890.
OS Wigglesworthia glossinidia brevipalpis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Wigglesworthia.
OX NCBI_TaxID=36870;
RN [1]
RC SEQUENCE FROM N.A.
RC MEDLINE=22297718; PubMed=12219091;
RA Akman L., Yamashita A., Watanabe H., Oshima K., Shiba T., Hattori M.,
RA Akao S.;
RT "Genome sequence of the endocellular obligate symbiont of tsetse
RT flies, Wigglesworthia glossinidia."
RL Nat. Genet. 32:402-407(2002).
CC -1- FUNCTION: Responsible for the release of ribosomes from messenger
CC RNA at the termination of protein biosynthesis. May increase the
CC efficiency of translation by recycling ribosomes from one round of
CC translation to another (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the RRF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB063522; BAC24535.1; -.
DR HAMAP; MF_00040; -; 1.
DR InterPro; IPR002661; RRF.
DR Pfam; PF01765; RRF; 1.
DR ProDom; PD004103; RRF; 1.
DR TIGRFAMs; TIGR00496; frr; 1.
KW Protein biosynthesis; Complete proteome.
SQ SEQUENCE 185 AA; 21650 MW; C71569E7155DECDF CRC64;

Query Match 54.4%; Score 37; DB 1; Length 185;
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 TELAKLVNKRSE 14
DB 354 TELAKLVNKRSE 365

RESULT 13
RRF WIGBR STANDARD; PRT; 185 AA.
AC Q8DZG5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ribosome recycling factor (ribosome releasing factor) (RRF).

```

```

QY 4 ELAKLVNKRSE 14
DB 112 ELAKLVNKRSE 122

RESULT 14
LACR STAM STANDARD; PRT; 251 AA.
AC Q99S73;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lactose phosphotransferase system repressor.
GN LACR OR SAV2196 OR SAV1998.
OS Staphylococcus aureus (strain Mu50 / ATCC 700639), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878; 158879;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700639, and N315;
RC MEDLINE=2111952; PubMed=11418146;
RA Kuroda M., Ohba T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Sekimizu K., Hirakawa H., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RT Lancet 357:1225-1240(2001).

```

```

CC -1- FUNCTION: Repressor of the lactose catabolism operon. Galactose-
CC 6-phosphate is the inducer.
CC -1- SIMILARITY: BELONGS TO THE DEOR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP003354; BAB58358.1; -
DR EMBL; AP003136; BAB43288.1; -
DR PIR; G90015; G90015.
DR InterPro; IPR001034; HTH_Deor.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00455; deor; 1.
DR PRINTS; PR00037; HTHLACR.
DR SMART; SM00420; HTH_Deor; 1.
DR PROSITE; PS00894; HTH_Deor_FAMILY; 1.
KW Lactose metabolism; Transcription regulation; Repressor; DNA-binding;
KW Complete proteome.
FT DNA_BIND 20
SQ SEQUENCE 251 AA; 28593 MW; C61ER3136AC0B37 CRC64;

Query Match 54.4%; Score 37; DB 1; Length 251;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 ELAKLVNKR 12
DB 9 ELAKLVNKK 17

RESULT 15
LACR_STAAM STANDARD; PRT; 251 AA.
ID LACR_STAAM STANDARD; PRT; 251 AA.
AC P16644;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lactose phosphotransferase system repressor.
GN LACR OR KW2122.
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620, 1280;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Ogunchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.,
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RK MEDLINE=90299802; PubMed=2163387;
RA Oskouian B., Stewart G.C.;
RT "Repression and catabolite repression of the lactose operon of
RT Staphylococcus aureus."
RL J. Bacteriol. 172:3804-3812(1990).
CC -1- FUNCTION: Repressor of the lactose catabolism operon. Galactose-
CC 6-phosphate is the inducer.
CC -1- SIMILARITY: BELONGS TO THE DEOR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP004829; BAB59987.1; -
DR EMBL; M22103; AAA67854.1; -
DR PIR; A44506; A44506.
DR InterPro; IPR001034; HTH_Deor.
DR Pfam; PF00455; deor; 1.
DR PRINTS; PR00037; HTHLACR.
DR SMART; SM00420; HTH_Deor; 1.
DR PROSITE; PS00894; HTH_Deor_FAMILY; 1.
KW Lactose metabolism; Transcription regulation; Repressor; DNA-binding;
KW Complete proteome.
FT DNA_BIND 20
SQ SEQUENCE 251 AA; 28549 MW; 9997A1DA3211AC4E CRC64;

Query Match 54.4%; Score 37; DB 1; Length 251;
Best Local Similarity 77.8%; Pred. No. 27;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 ELAKLVNKR 12
DB 9 ELAKLVNKK 17

Search completed: May 21, 2004, 18:02:16
Job time : 13 secs

```



```

051873      PRELIMINARY;      PRT;      314 AA.
ID 051873;
AC 051873;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Proline dehydrogenase (Fragment).
GN PUTA.
OS Photobacterium leiognathi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=658;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PL741;
RX MEDLINE=96216743; PubMed=8645272;
RA Lin J.W., Yu K.Y., Chen H.Y., Weng S.F.,
RT "Regulatory region with puta gene of proline dehydrogenase that links
RT to the lum and the lux operons in Photobacterium leiognathi."
RL Biochem. Biophys. Res. Commun. 219:868-875(1996).
DR EMBL: U99227; AAC43866.1; -.
DR PIR; JC4661; JC4661.
DR GO; GO:0004657; F:proline dehydrogenase activity; IEA.
DR GO; GO:0006537; P:glutamate biosynthesis; IEA.
DR GO; GO:0006562; P:proline catabolism; IEA.
DR InterPro: IPR002872; Pro_dh.
DR Pfam: PF01619; Pro_dh; 1.
FT NON TER 314 314
SQ SEQUENCE 314 AA; 35213 MW; 6E85DBC8E8225B7 CRC64;

Query Match      64.7%; Score 44; DB 2; Length 314;
Best Local Similarity 64.3%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 TPTLAKLVNKRSE 14
Db      154 TPTVIRLVNKRSE 167

RESULT 3
Q9UN52      PRELIMINARY;      PRT;      423 AA.
ID 09UN52;
AC 09UN52; 043191;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE COP9 complex subunit 3 (COP9 (Constitutive PHOTOMORPHOGENIC,
DE ARABIDOPSIS, homolog) subunit 3) (YAB1-containing signalosome subunit
DE 3).
GN COP9 OR SGN3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99208679; PubMed=10191102;
RA Potocsi L., Chen K.-S., Lupski J.R.;
RT "Subunit 3 of the COP9 signal transduction complex is conserved from
RT plants to humans and maps within the smith-magenis syndrome critical
RT region in 17p11.2."
RT Genomics 57:180-182(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX Strauberg R.;
RA submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 21-423 FROM N.A.
RX MEDLINE=98194867; PubMed=9535219;
RA Seeger M., Kraft R., Ferrell K., Bech-Oschar D., Dundey R.,
RA Schade R., Gordon C., Naumann M., Dubiel W.;

```

```

RT "A novel protein complex involved in signal transduction possessing
RT similarities to 26S proteasome subunits."
RL PAPER J. 12:469-478(1998)
DR EMBL: AF098109; AAD41247.1; -.
DR EMBL: BC001891; AAH01891.1; -.
DR EMBL: AF031647; AAC14197.1; -.
DR Genew; HGNC:2239; COPS3.
DR GO; GO:0005737; Cytoplasm; TAS.
DR GO; GO:0009416; P:response to light; TAS.
DR GO; GO:007165; P:signal transduction; TAS.
DR InterPro: IPR000717; PCT.
DR Pfam; PF01399; PCT; 1.
DR SMART; SM00086; PINT; 1.
SQ SEQUENCE 423 AA; 47873 MW; 1D371050C7D7BFB8D CRC64;

Query Match      64.7%; Score 44; DB 4; Length 423;
Best Local Similarity 69.2%; Pred. No. 17;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 PTEIAKLVNKRSE 14
Db      272 PSEIRLVNKRSE 284

RESULT 4
Q89E44      PRELIMINARY;      PRT;      449 AA.
ID 089E44;
AC 089E44;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BLR7243 protein.
GN BLR7243.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Ideawa K., Iriyuchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005961; BAC52508.1; -.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Complete proteome.
SQ SEQUENCE 449 AA; 51846 MW; 6FB122A36F2CC0F CRC64;

Query Match      62.5%; Score 42.5; DB 16; Length 449;
Best Local Similarity 58.8%; Pred. No. 33;
Matches 10; Conservative 3; Mismatches 1; Indels 3; Gaps 1;

QY      1 TPT---ELAKLVNKRSE 14
Db      154 TPTHLDELAKLVNKRSE 170

RESULT 5
Q7ZVT8      PRELIMINARY;      PRT;      423 AA.
ID 07ZVT8;
AC 07ZVT8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to COP9 constitutive photomorphogenic homolog subunit 3

```

```

DE (Arbidiopsis).
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045415; AA045415.1; -.
DR InterPro; IPR000717; PCI.
DR InterPro; IPR008941; TPR-1like.
DR Pfam; PF01359; PCI; 1.
DR SMART; SM00088; PINT; 1.
SQ SEQUENCE 423 AA; 47909 MW; 88D075657C0F7C65 CRC64;

Query Match 61.8%; Score 42; DB 13; Length 423;
Best Local Similarity 69.2%; Pred. No. 39;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 PTELAKLVNKRSE 14
DB 272 PRLRALVYNCHSE 264

RESULT 6
Q8YZU6 PRELIMINARY; PRT; 248 AA.
AC Q8YZU6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Hypothetical protein Alr0359.
GN Alr0359.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasaamoto S.,
RA Kishida A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003582; BAB72317.1; -.
DR PIR; AF1851; AF1851.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 248 AA; 27298 MW; 0AE8145DD0D05991 CRC64;

Query Match 60.3%; Score 41; DB 16; Length 248;
Best Local Similarity 72.7%; Pred. No. 34;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PTELAKLVNKR 12
DB 225 PTELKLVNKR 235

RESULT 7
Q8D910 PRELIMINARY; PRT; 399 AA.
AC Q8D910;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Uncharacterized conserved protein.
GN VY12621.

```

```

OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCPC6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCPC6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016805; AA010970.1; -.
KM Complete proteome.
SQ SEQUENCE 399 AA; 43925 MW; E2F5F951FA95E132 CRC64;

Query Match 60.3%; Score 41; DB 16; Length 399;
Best Local Similarity 64.3%; Pred. No. 55;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PTELAKLVNKRSE 14
DB 204 PTELISGLVTKVSD 217

RESULT 8
Q87C79 PRELIMINARY; PRT; 135 AA.
ID Q87C79;
AC Q87C79;
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Conserved hypothetical protein.
GN P1214.
OS Xylella fastidiosa (strain Temecula / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
RA Miyaki C.Y., Furlan L.R., Camargo L.B.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
RA Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
RA Coutinho L.B., Kimura E.T., Ferro E.S., Harakava R., Kuriama E.E.,
RA Martino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Penille R.C., Ferro J.A., Fontighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Teukuno F., Yanai G.M., Zeros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa."
RL J. Bacteriol. 185:1018-1026(2003).
DR EMBL; AE012557; AA029065.1; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 135 AA; 15769 MW; 7C54096E7B99455 CRC64;

Query Match 58.8%; Score 40; DB 16; Length 135;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 TELAKLVNKRSE 14
DB 28 TDVAKLVNKRNE 39

RESULT 9
Q7ML59 PRELIMINARY; PRT; 162 AA.
ID Q7ML59

```

```

AC 07W59;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator (Marr family).
GN BB1890.
OS Bordetella bronchiseptica (Alcaligenes bronchiseptus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaltha M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsis K.,
RA Leather S., Moulé S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabbinnovitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640442; CAB2387.1; -.
KM Complete proteome.
SQ SEQUENCE 162 AA; 18047 MW; 96814512C9198346 CRC64;

Query Match 58.8%; Score 40; DB 16; Length 162;
Best Local Similarity 70.0%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPELAKLVN 10
Db 69 TPAELARLIN 78

RESULT 10
07W52 PRELIMINARY; PRT; 162 AA.
ID 07W52;
AC 07W52;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator (Marr family).
GN BPP2442.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaltha M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsis K.,
RA Leather S., Moulé S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabbinnovitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640442; CAB2387.1; -.
KM Complete proteome.
SQ SEQUENCE 162 AA; 18047 MW; 96814512C9198346 CRC64;

```

```

Query Match 58.8%; Score 40; DB 16; Length 162;
Best Local Similarity 70.0%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPELAKLVN 10
Db 69 TPAELARLIN 78

RESULT 11
07W58 PRELIMINARY; PRT; 162 AA.
ID 07W58;
AC 07W58;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Putative transcriptional regulator (Marr family).
GN BP2286.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebaltha M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsis K.,
RA Leather S., Moulé S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabbinnovitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640417; CAB42560.1; -.
KM Complete proteome.
SQ SEQUENCE 162 AA; 18047 MW; 96814512C9198346 CRC64;

Query Match 58.8%; Score 40; DB 16; Length 162;
Best Local Similarity 70.0%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TPELAKLVN 10
Db 69 TPAELARLIN 78

RESULT 12
087440 PRELIMINARY; PRT; 390 AA.
ID 087440;
AC 087440;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroviruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=9513252; PubMed=786956;
RX "Progression to AIDS in macaques is associated with changes in the
RT replication, tropism, and cytopathic properties of the simian
RT immunodeficiency virus variant population.";
RL Virology 207:528-542(1995).

```

DR EMBL: U09092; AAA73908.1; -
 DR GO: GO:0019028; C:Viral capsid; IEA.
 DR GO: GO:0019031; C:Viral envelope; IEA.
 DR GO: GO:0005198; F:Structural molecule activity; IEA.
 DR InterPro: IPR000777; GP120.
 DR Pfam: PF00516; GP120; 1.
 KM AIDS; Coat protein; Glycoprotein.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 390 AA; 45044 MW; 19AD4C3A13F81532 CRC64;
 Query Match 58.8%; Score 40; DB 15; Length 390;
 Best Local Similarity 61.5%; Pred. No. 82;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 TPTEALVKNRS 13
 DB 36 TPTEALVKNRS 48
 RESULT 13
 Q8YFH4 PRELIMINARY; PRT; 392 AA.
 AC Q8YFH4;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE Spore-cortex-lytic enzyme PREPPRIIDE precursor.
 GN BME11548.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 NCBI_TaxID=29459;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=1175668;
 RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Leveson J.-J.,
 RA Haselkorn R., Kyrides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
 DR EMBL: AEO09590; AAL52729.1; -
 DR PIR: AF3445; AF3445.
 KM Complete proteome.
 SQ SEQUENCE 392 AA; 43099 MW; B8BDA0AFA3C8C06 CRC64;
 Query Match 58.8%; Score 40; DB 16; Length 392;
 Best Local Similarity 77.8%; Pred. No. 83;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 PTEALVKN 10
 DB 190 PTEALVKN 198
 RESULT 14
 Q8G2E5 PRELIMINARY; PRT; 429 AA.
 AC Q8G2E5;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMblrel. 23, Last annotation update)
 DE Conserved hypothetical protein.
 GN BR0378.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 NCBI_TaxID=29461;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=130 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Bisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Brennan M.J.,
 RA Daugherty S.C., Debay R.T., Durkin A.S., Kolonay J.P., Madupu R.,
 RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek U., Van Aken S.E.,
 RA Riedmiller S., Tetteil H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
 DR EMBL: AEO14349; AAN29324.1; -
 DR TIGR: BR0378;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 429 AA; 47131 MW; 56F3752AEA471DID CRC64;
 Query Match 58.8%; Score 40; DB 16; Length 429;
 Best Local Similarity 77.8%; Pred. No. 91;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 PTEALVKN 10
 DB 227 PTEALVKN 235
 RESULT 15
 Q9FRL7 PRELIMINARY; PRT; 593 AA.
 ID Q9FRL7
 AC Q9FRL7;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN P22H5.2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
 RA Barnstead M.E., Bowman C.L., White O., Niernm W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome 1 BAC P22H5 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC025814; AAG12682.1; -
 DR PIR: A96783; A96783.
 DR GO: GO:0003824; F:catalytic activity; IEA.
 DR InterPro: IPR003865; Isoflav reduct.
 DR Pfam: PF02716; Isoflavone_redu; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 593 AA; 65710 MW; 9840111808C0F0E CRC64;
 Query Match 58.8%; Score 40; DB 10; Length 593;
 Best Local Similarity 61.5%; Pred. No. 1.3e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 TPTEALVKNRS 13
 DB 85 TPTEALVKNRS 97
 Search completed: May 21, 2004, 18:03:09
 Job time : 42 secs